Copernicus
Europe's eyes on Earth

Copernicus Atmosphere Monitoring Service

Validation report of the CAMS near-real time global atmospheric composition service

June – August 2019

European Commission

Copernicus
Europe's eyes on Earth

IMPLEMENTED BY
ECMWF

Validation report of the CAMS near-real-time global atmospheric composition service. Period June - August 2019

EDITORS:
M. Ramonet (LSCE), A. Wagner (MPG), M. Schulz (MetNo),
Y. Christophe (BIRA-IASB), H.J. Eskes (KNMI)

AUTHORS:
S. Basart (BSC), A. Benedictow (MetNo), Y. Bennouna (CNRS-LA),
A.-M. Blechschmidt (IUP-UB), S. Chabrillat (BIRA-IASB), E. Cuevas (AEMET),
A. El-Yazidi (LSCE), H. Flentje (DWD), K. M. Hansen (AU), U. Im (AU),
J. Kapsomenakis (AA), B. Langerock (BIRA-IASB), A. Richter (IUP-UB),
N. Sudarchikova (MPG), V. Thouret (CNRS-LA), T. Warneke (UBC),
C. Zerefos (AA)

REPORT OF THE COPERNICUS ATMOSPHERE MONITORING SERVICE, VALIDATION SUBPROJECT.

Executive Summary

The Copernicus Atmosphere Monitoring Service (http://atmosphere.copernicus.eu, CAMS) is a component of the European Earth Observation programme Copernicus. The CAMS global near-real time (NRT) service provides daily analyses and forecasts of reactive trace gases, greenhouse gases and aerosol concentrations. This document presents the validation statistics and system evolution of the CAMS NRT service for the period up to 1 September 2019, with a focus on June-August 2019 (JJA-2019). Updates of this document appear every 3 months, e.g. Christophe et al. (2019).

This summary is split according to service themes as introduced on the CAMS website: air quality & atmospheric composition, climate forcing, ozone layer and UV. Specific attention is given to the ability of the CAMS system to capture recent events. We focus on the 'o-suite' composition fields, which are the daily analyses and forecasts produced by the IFS (Integrated Forecast System) modelling system at ECMWF, using the available meteorological and atmospheric composition observations which are ingested in the ECMWF 4D-Var assimilation system. The model and assimilation configurations are summarised in section 2. We furthermore assess the impact of the composition observations by comparing the validation results from the 'o-suite' to a 'control' configuration without atmospheric composition data assimilation. Also, the pre-operational delayed-mode analyses and high-resolution forecasts of CO_2 and CH_4 are assessed in this report.

A detailed description of the measurement datasets used is provided in Eskes et al. (2018a).

On 9 July 2019, in the middle of the reporting period of this report, a major upgrade of the CAMS system to version 46R1 took place. Among other things this involved a change from 60 vertical levels to 137 vertical levels. The upgrade is described in more detail in section 2.

The o-suite data delivery for the period JJA-2019 was very good, with an on-time percentage of 99.46 %. Only on 21 August there was a 7 min delay with some of the products.

Air quality and atmospheric composition

Tropospheric ozone (O₃)

The CAMS o-suite ozone is validated with surface and free tropospheric ozone observations from the WMO Global Atmospheric Watch (GAW) and NOAA Earth System Research Laboratory (ESRL) networks, In-service Aircraft for a Global Observing System (IAGOS) airborne data, ozonesondes, and the MetOp Infrared Atmospheric Sounding Interferometer (IASI) tropospheric ozone retrievals. For column-averaged free tropospheric ozone against ozonesondes the o-suite modified normalized mean biases (MNMBs) are on average small, ±20% for all months over the Arctic and Northern Midlatitudes (Fig. S.1). Over Antarctica and the Tropics o-suite biases are up to +30% for recent years. The control run mostly shows negative biases.

For IAGOS comparisons, no notable difference is found in the ozone profiles when comparing the period before and after the upgrade to CY46R1 in July 2019. The comparison is complicated by a measurement issue with one of the IAGOS instruments. During the heat wave periods over Europe, the models reproduce the related ozone increase at Frankfurt but underestimate the peak values
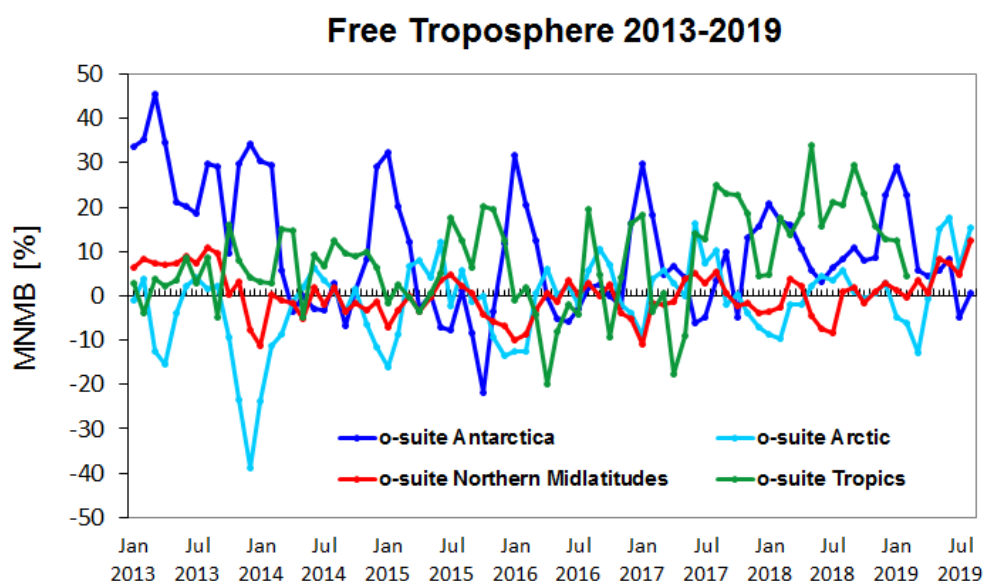
Free Troposphere 2013-2019

Figure S.1: Time series of MNMB of ozone in the o-suite, compared against ozonesondes, averaged over different latitude bands, for the period January 2013 to August 2019. The free troposphere is defined here as the layer between 750 and 300 hPa.

with slightly better performance from the o-suite as compared to the control run. In the UTLS region, ozone is generally overestimated by the o-suite and partly better represented by the control run, in particular over North America and the Middle East. Good agreement with the IAGOS free troposphere ozone profiles is also found over the other regions of the world.

In comparison with surface observations we find a steady improvement of the o-suite over the past 5 years over European GAW stations. Biases are around ±10% for recent years, and within ±20% (the Arctic is discussed below). Mean MNMBs of European stations are 6% for surface ozone for Europe during June to August 2019. Both runs overestimate O_3 observations for Asia with MNMBs up to 50% for the o-suite and the control run. For the tropics, surface ozone is overestimated with MNMBs within 15%. Both runs can reproduce Antarctic surface observations with MNMBs 25%.

Tropospheric Nitrogen dioxide (NO₂)

Model validation against satellite NO_2 measurements, using Envisat SCIAMACHY data before April 2012 and MetOp-A GOME-2 data afterwards, shows that tropospheric NO_2 columns are well reproduced by the NRT model runs, indicating that emission patterns and NO_x photochemistry are generally well represented, although modelled shipping signals are more pronounced than in the satellite retrievals. Tropospheric NO_2 columns over some local emission hotspots (e.g. Moscow, and Red Basin in China) are overestimated, while wintertime and springtime values over Europe around Benelux are underestimated. Since December 2014, the agreement between satellite retrievals and model results for time series over East-Asia and Europe is better than for previous years (Fig. S.2), as observed columns of NO_2 decreased in 2014, likely associated with reduced emissions, and (in contrast to the observations) simulated values show an increase over the whole timeseries available. Between spring and autumn, the models regularly show an overestimation over several regions with boreal forest fire activity (Canada, Alaska, Siberia).
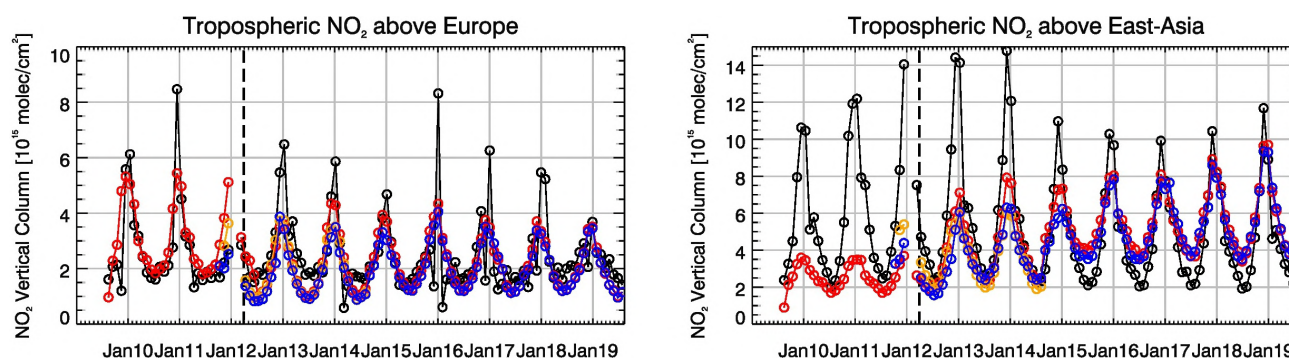
Figure S.2: Time series of tropospheric NO_2 columns from SCIAMACHY (up to March 2012) and GOME-2 (from April 2012 onwards) compared to MACC/CAMS model results for Europe and East-Asia. The o-suite is in red, control is in blue (the model run without data assimilation is termed control since Sep 2014).

Tropospheric Carbon Monoxide (CO)

The CAMS products are validated with GAW network surface observations, IAGOS airborne data, Fourier Transform InfraRed (FTIR) observations from the Network for the Detection of Atmospheric Composition Change (NDACC) and the Total Carbon Column Observing Network (TCCON), and Measurements of Pollution In The Troposphere (MOPITT) and IASI satellite CO retrievals. The comparisons reveals that the absolute values, latitude dependence and seasonality, as well as day-to-day variability of CO can be reproduced well by the CAMS-global analyses and forecasts. Biases are between -6% and 5% for European GAW stations, and up to -13% in Asia.

For stations in the southern hemisphere the comparison with NDACC (Fig. S.3) and GAW measurements show, that data assimilation efficiently reduces the large positive MNMBs observed in the control run. The o-suite compared with the TCCON CO observations shows basically no latitude dependence. Overall, a small positive bias of < 5% is found with TCCON. Similar results are obtained when comparing to the NDACC FTIR measurements, but now with a small negative bias of a few % for the tropospheric CO column. A small seasonal cycle is observed in the biases, but this stays within 5%.

According to IAGOS aircraft observations CO is mostly underestimated over Frankfurt by both the o-suite and control run, and the largest bias is found in the lowest layers. The performance of the two runs is similar in the lowest layers, while in the free troposphere the performance of the o-suite is better than that of control run. However, the bias in CO appears improved after the upgrade to CY46R1 in July 2019 especially in the lowest layers while for control run no change is noted. For most other regions of the world the results of the two CAMS configurations are similar as for Europe.

The comparisons with MOPITT and IASI confirm these findings. The day 0 o-suite forecast shows differences within 10% regionally compared to MOPITT (-5% to -10% for Europe), with little latitude dependence of the bias. Regionally these biases increase during the 4-day forecast (larger positive biases over Siberia, Canada and Central Africa, larger negative biases in some locations over the continents). A general reduction of CO concentrations from the year 2015 to the year 2018 can be seen over Europe, the US and East Asian regions. Due to the stability of the (small) bias this trend is well reproduced by the o-suite.
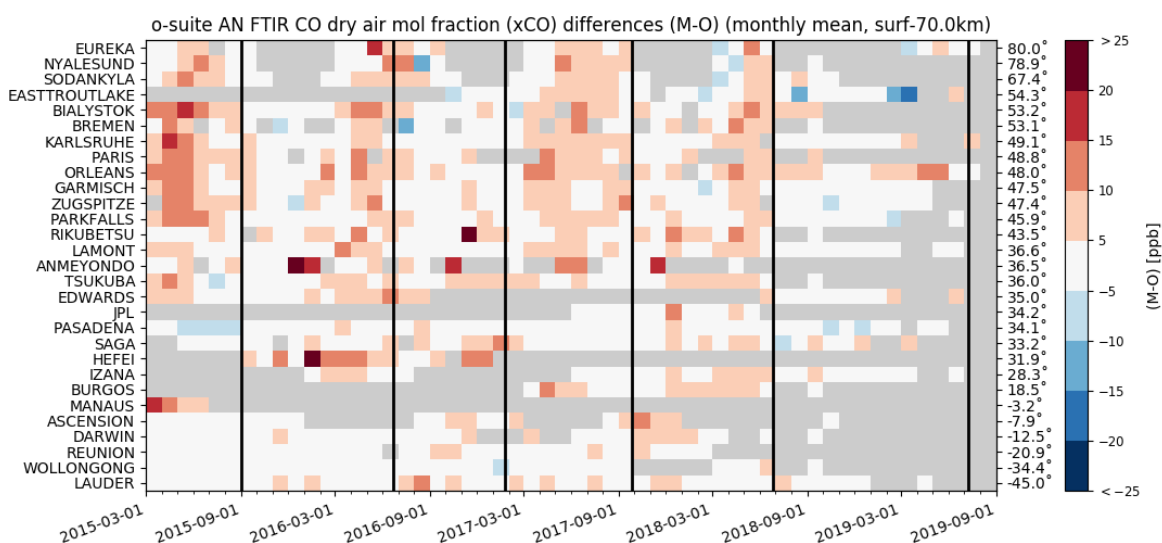
Figure S.3: Monthly mean relative CO bias (o-suite – observation)/observation for the last 4 years. Comparisons are made against TCCON FTIR CO remote sensing observations. Model upgrades are indicated with the black lines. The stations are sorted with decreasing latitude (northern to southern hemisphere). A similar plot for the control run shows biases up to 40% in the southern hemisphere.

Formaldehyde

Model validation, with respect to SCIAMACHY/Envisat HCHO data before April 2012 and GOME-2/MetOp-A HCHO data afterwards (Fig. S.4), shows that modelled monthly HCHO columns represent well the magnitude of oceanic and continental background values and the overall spatial distribution in comparison with mean satellite HCHO columns. Compared to GOME-2 satellite retrievals, an overestimation of values regularly occurs over Australia and Central Africa, which could be both related to biogenic emissions and fire emissions. For time series over East-Asia and the Eastern US, both regions where HCHO columns are probably dominated by biogenic emissions, models and retrievals agree rather well, but the yearly cycle over East-Asia is underestimated by the models.

Aerosol

We estimate that the o-suite aerosol optical depth showed an average positive bias in the latest three months of +18%, measured as modified normalized mean bias against daily Aerosol Robotic NETwork (Aeronet v3 level 1.5) sun photometer data. The 3-day forecasted aerosol distribution shows 19% less aerosol optical depth (AOD) than that from the initial forecast day, as shown in Fig. S.5-a. The spatiotemporal correlation, shown in Fig. S.5-b, shows month-to-month variation in JJA 2019 similar to summer 2018, indicating the simulation reproduces approximately 64% of the day to day AOD variability across all Aeronet stations. The o-suite forecast at +3 days shows slightly lower correlation, as a consequence of imperfect forecasted meteorology and fading impact of the initial assimilation of MODIS AOD and MODIS fire info on model performance. The o-suite forecast running each day at 12UTC shows almost identical performance as the forecast starting at 00UTC.
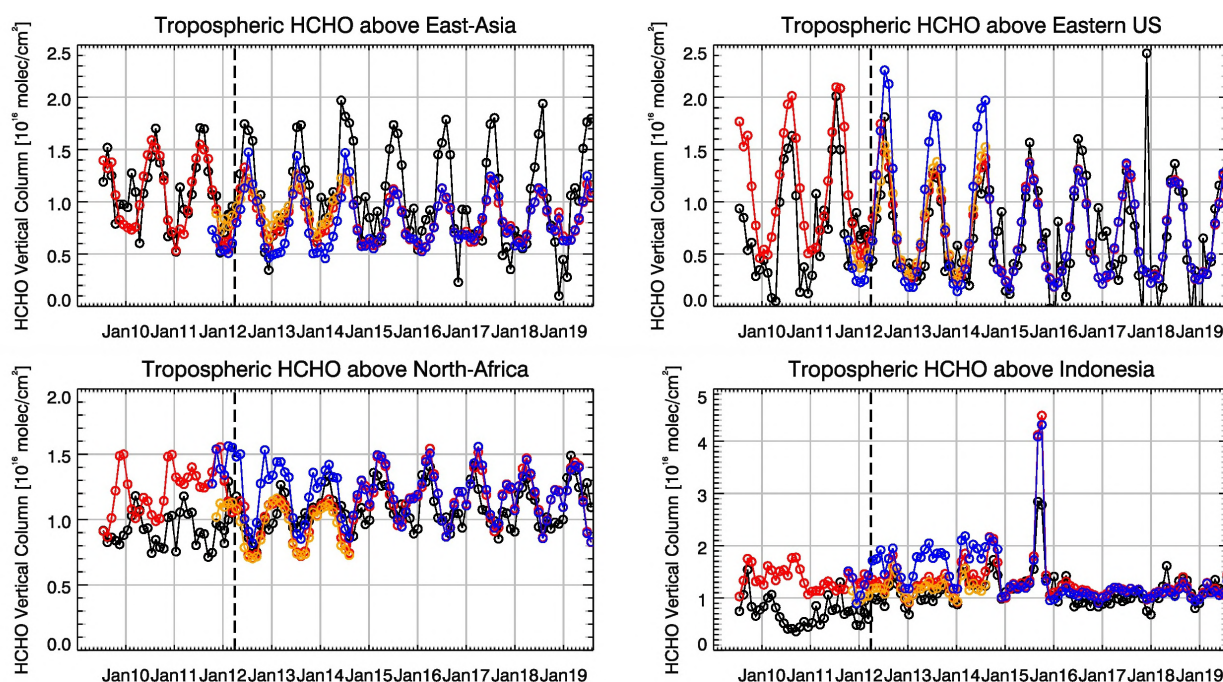
Figure S.4: Time series of average tropospheric HCHO columns [10^{16} molec cm^{-2}] from SCIAMACHY (up to March 2012) and GOME-2 (from April 2012 onwards) compared to model results for different regions. The blue line shows CAMS control results including older configurations from the MACC project before September 2014. The regions are: East-Asia (25-40°N, 110-125°E), Eastern US (30-40°N, 75-90°W), Northern Africa (0-15°N, 15°W-25°E) and Indonesia (5°S-5°N, 100-120°E). Vertical dashed black lines mark the change from SCIAMACHY to GOME-2 based comparisons in April 2012.

The AOD performance of the o-suite with respect to the AERONET data exhibits no pronounced seasonal cycle. Correlations are smaller in late summer. Since October 2017, the largest contributions to global AOD come from organics and sea salt. Sea salt AOD increased further, especially in the tropics, due to the model upgrade to CY45R1 with the new sea salt emission scheme activated, while dust AOD became lower compared to earlier years. With the coupling of chemistry and aerosol schemes for sulphur in the latest upgrade to CY46R1 in July 2019, there is an increase of SO$_4$ especially in the northern hemisphere.

The aerosol Ångström exponent (AE) contains information about the size distribution of the aerosol, and implicitly about composition. In the last year the o-suite AE became more positive indicating a change to slightly more fine particles since the model upgrade to version 45R1 in June 2018.

PM10 and PM25, as defined by the IFS aerosol model, are evaluated against an average from rural and background site data in the period 2000-2009 at 160 sites in North America and Europe. This indicates that PM10 concentrations exhibit on average in the latest period an underestimation with MNMB bias of -10% in Europe and of -1% in North America. PM25 concentrations are overestimated 3% in Europe and 38% in North America. Consistent with this finding a higher positive bias is also found for AOD in North America than in Europe. The fraction of PM simulated data within a factor 2 of observed values stayed similar since September 2017 for both PM10 and PM25. PM25 seems to have deteriorated compared to periods before mid 2017, while PM10 shows an improvement. However, with the model version upgrade to CY46R1 in July 2019 it seems that PM25 has improved significantly.
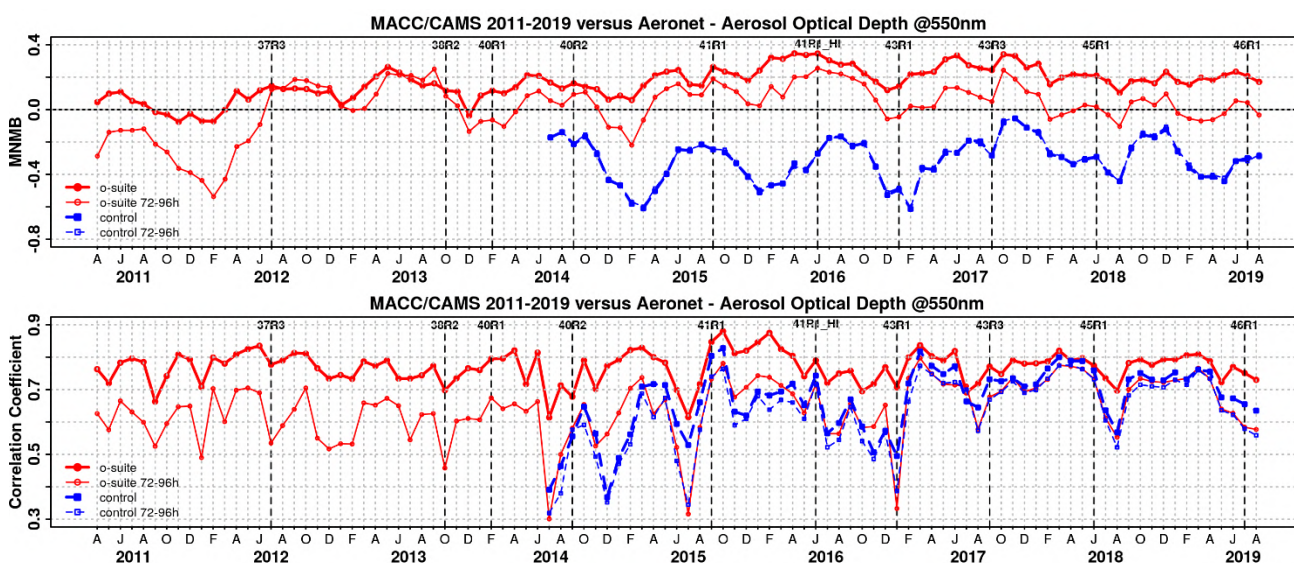
Figure S.5. Aerosol optical depth at 550nm in IFS 00Z model simulations for April 2011 – August 2019 against daily matching Aeronet Version3 level 1.5 data. a) Modified normalized mean bias (MNMB); o-suite (thick red curve); o-suite at last forecast day (light red curve); Control (blue dashed); Control at last forecast day (light blue dashed); b) Corresponding correlation coefficient. Model version changes are marked as vertical bars.

During this season, satellites show that major dust activity in Northern Africa (seasonal AOD up to 0.3) is concentrated in latitudes south of 20ºN with maximum seasonal values (seasonal AOD over 0.7). In comparison with MODIS (see Figure S.6), both CAMS experiments underestimate dust activity over the Bodélé (in Chad) and overestimate dust observations in the source region between Mali-Algeria-Niger borders as well as Sudan, Iraq and North Saudi Arabia. The CAMS o-suite presents lower season values (seasonal DOD up to 0.9) than control (seasonal DOD up to 1.2), which are both in general higher than the WMO Sand and Dust Storm Warning Advisory and Assessment System (SDS-WAS) multi-median product (seasonal DOD up to 0.5). CAMS o-suite can reproduce the dust transport over the North Atlantic region although the maximum dust activity is shifted to Mali, Niger and Algeria border and Eastern Sahara instead in Chad as it is shown in MODIS. Also, DOD over Iraq and the Mediterranean Basin appears overestimated in the comparison with MODIS.

From June to August, the o-suite reproduces the daily variability of AERONET direct-sun observations with a correlation coefficient of 0.82, averaged over all the AERONET sites (as in the case of the SDS-WAS multi-model product), which is close to the control experiment that shows a correlation coefficient of 0.83. Regarding mean bias (MB), both CAMS experiments (o-suite and control) slightly overestimate the AERONET observations resulting in an MB of 0 for o-suite and 0.05 for control. Similar results are obtained in the comparison with the AERONET Spectral De-Convolution Algorithm retrievals.

The behaviour of the CAMS model changes at sites in desert regions after the upgrade to CY46R1 in July 2019. This upgrade included the implementation of a new dust module in the operational CAMS model and led to higher dust values in the CAMS control run at these sites. This overestimation is reduced by the assimilation in the o-suite.
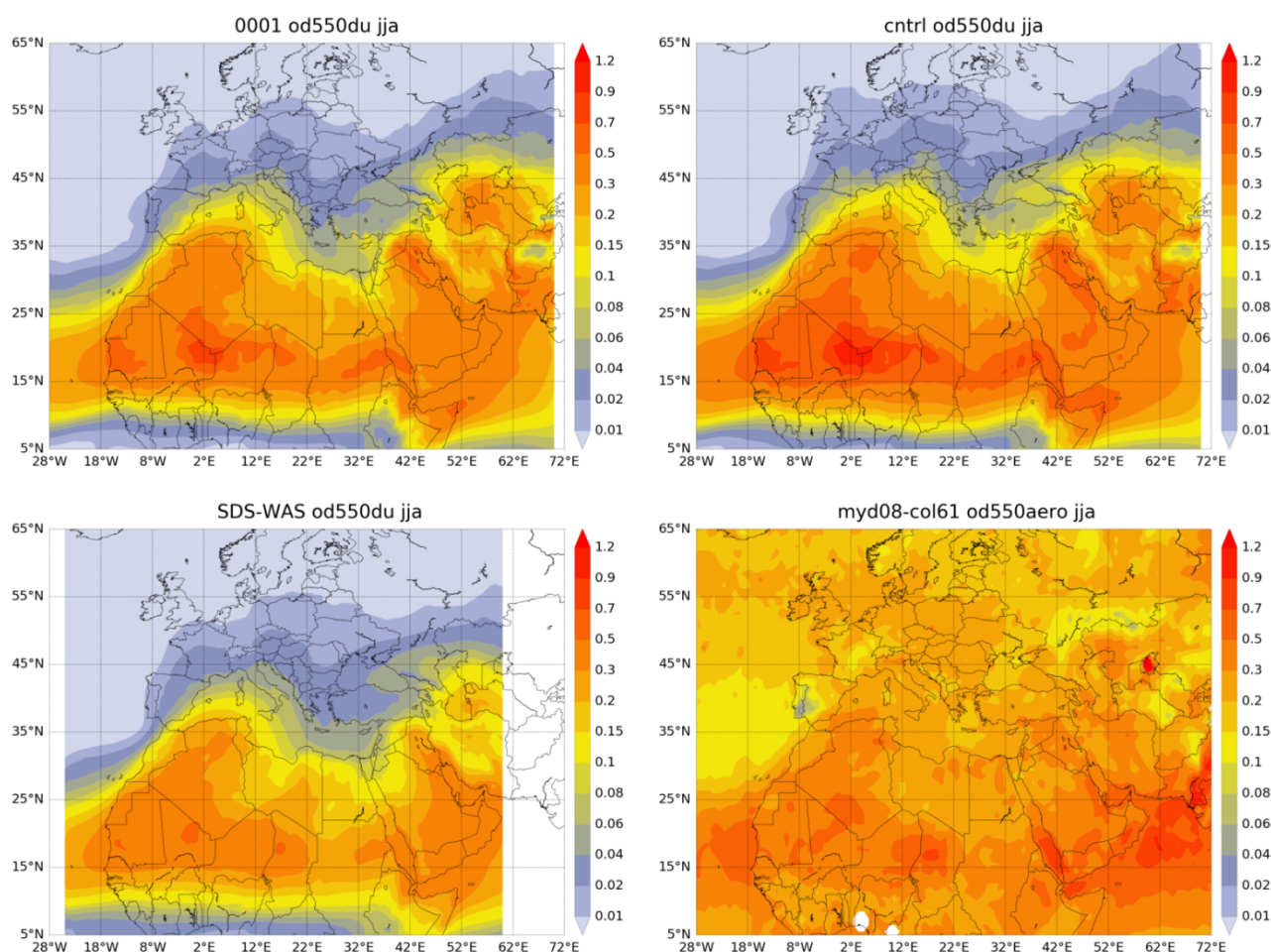
Figure S.6: Averaged DOD 24h forecast from o-suite (top left) and control (top right), DOD of the multi-model SDS-WAS Median product (bottom left) as well as AOD from MODIS/Aqua Collection 6.1 Level 3 combined Dark target and Deep Blue product (bottom right) for the study period.

The comparison of the 1- to 3-day forecasts shows that the forecast is stable during the 3-days forecasts in the comparison with AERONET direct-sun observations with correlation coefficients of 0.82 (0.83), 0.81 (0.83), and 0.79 (0.81) respectively for 24, 48 and 72h forecasts for all the sites for o-suite (control).

The comparison with the German ceilometer network shows that the general low bias of model aerosol backscatter in the planetary boundary layer has disappeared or is turned into a positive bias after the July 2019 upgrade to CY46R1. This may be related to the addition of nitrate and ammonia in the model since July 2019. The high bias of the model backscatter coefficients in the free troposphere has not changed notably. The profile is still strongly smoothed, i.e. the observed step in concentration at the top of the boundary layer is not captured notably better with 137 levels than with 60 levels (51 layers instead of 27 layers below 8 km altitude), and this conclusion is the same for o-suite and control.

System performance in the Arctic

The CAMS model runs are validated using surface ozone measurements from the ESRL-GMD and the International Arctic Systems for Observing the Atmosphere (IASOA) network (5 sites) and ozone concentrations in the free troposphere and UTLS are evaluated using balloon sonde measurement data.

From June to August 2019 the simulations of the surface ozone concentrations are on average in good agreement with the observations with MNMB between -14% and 11%.

During June – August 2019 there is an overestimation of ozone concentrations in the Arctic free troposphere for the o-suite (MNMB = 6% – 18%) and for the control run (MNMB = -2% – 7%) as well as in the UTLS (MNMB up to 20% for the o-suite). The larger positive biases coincide with the stopping of the assimilation of OMPS data from May 2019 onwards.

Total column O_3 is in good agreement with the observations obtained from IASI, showing MNMBs within 5%. Note that the IASI sensitivity is the lowest over the cold surfaces such as the Greenland ice sheet where IASI O_3 values are positively biased by up to 20%.

Comparison with FTIR observations from the NDACC network shows that the CO tropospheric columns are in good agreements at the arctic sites with bias between -1% and -7%, with larger negative bias for the the control run (-9 – -14%). The bias of the stratospheric CO column is improved with the model upgrade (60 to 137 levels) implemented in July 2019. The total column CH4 concentration is within ±2% at the three Arctic site, while the tropospheric column is slightly underestimated (-2 – -4%) at one of the sites.

Comparison with MOPITT version 8 shows that modelled CO total columns are in relative good agreement with the satellite retrievals with low bias in the Arctic (±20%), with a tendency for an underestimation over land and an overestimation over the ocean.

System performance in the Mediterranean

The CAMS o-suite reproduces the daily variability of AERONET direct-sun observations. In the Western, Central and Eastern Mediterranean, the correlation coefficient decreases from 0.82, 0.86 and 0.80, to 0.74, 0.82 and 0.65, respectively for control and o-suite during summer. Both CAMS experiments overestimated the AERONET observations in the Mediterranean Basin in control (MB of 0.11, 0.17 and 0.19 for Western, Central and Eastern Mediterranean regions respectively) and o-suite (MB of 0.13, 0.16 and 0.19 for Western, Central and Eastern Mediterranean regions respectively). The highest peaks on CAMS AOD simulations are linked to desert dust intrusions occurring during the whole season in the whole Mediterranean Basin.

For summer, PM10 and PM2.5 results of CAMS o-suite and control show similar skill scores in comparison with the European Environment Information and Observation Network (EIONET-Airbase) observations. The CAMS model tends to overestimate the PM10 and PM2.5 EIONET-Airbase observations in Central Europe with MB up to 10µg/m^3 while the PM10 and PM2.5 observed values are underestimated at the Iberian Peninsula and North Atlantic sites. Overall, for all the EIONET-Airbase sites, the o-suite presents higher overestimations in PM10 and PM2.5 (with MB of 2.20 and 3.45 µg/m^3, respectively) than control (with MB of -1.11 and 1.11 µg/m^3, respectively).
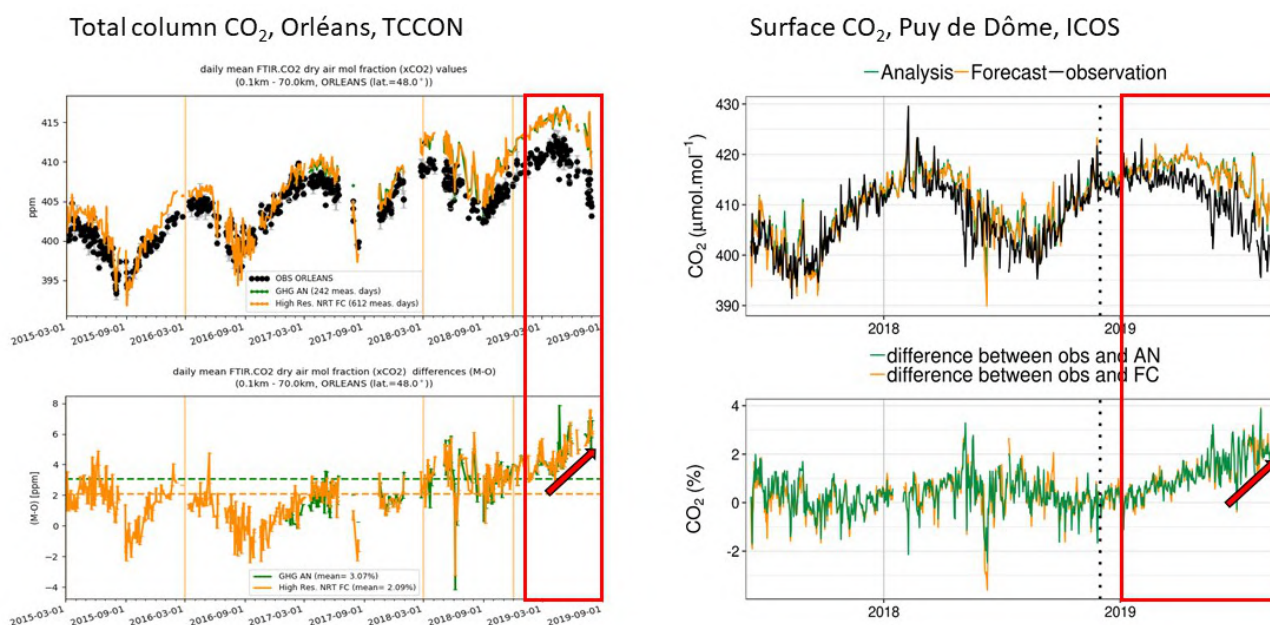
Figure S.7: Comparison of CO_2 measurements with simulations at the surface (right) in Puy de Dôme (France), and in the total column (left) at Orléans (France).

The upgrade of the CAMS model is clearly identified in the EIONET-Airbase time series. From 9th July, the PM10/PM2.5 ratio is higher than in previous periods indicating an increase of the coarse particles at surface levels, which introduces overestimations in the PM10 fractions.

Climate forcing

Greenhouse gases

CO_2 and CH_4 surface concentrations from the European Integrated Carbon Observation System (ICOS) network, and total or partial columns from TCCON and NDACC stations have been used to validate the analysis and high-resolution forecast experiments.

According to ICOS stations, the bias on CH_4 surface measurements is clearly dependent on latitude, at least in Europe. We observe a positive bias (20 to 50 ppb) at high latitude, and a negative one in Southern Europe (-10 to -30 ppb). Column measurements (TCCON and NDACC) indicate negative biases, but also appear to be dependent on altitude since NDACC measurements in the stratosphere show a slight overestimation compared to the measurement uncertainty.

The surface and total column measurements indicate an overestimation of the amplitude of the CO_2 seasonal cycle in the northern hemisphere by ±1%. The drought anomaly in spring/summer 2018 has an additional effect on the comparison with surface sites from May to July 2018, with an overestimated impact of the drought on the CO_2 concentrations. Both surface and total column measurements indicate an increase of the positive bias in the recent months (July-August 2019), at least in Europe, up to +6 ppm (Figure S.7).

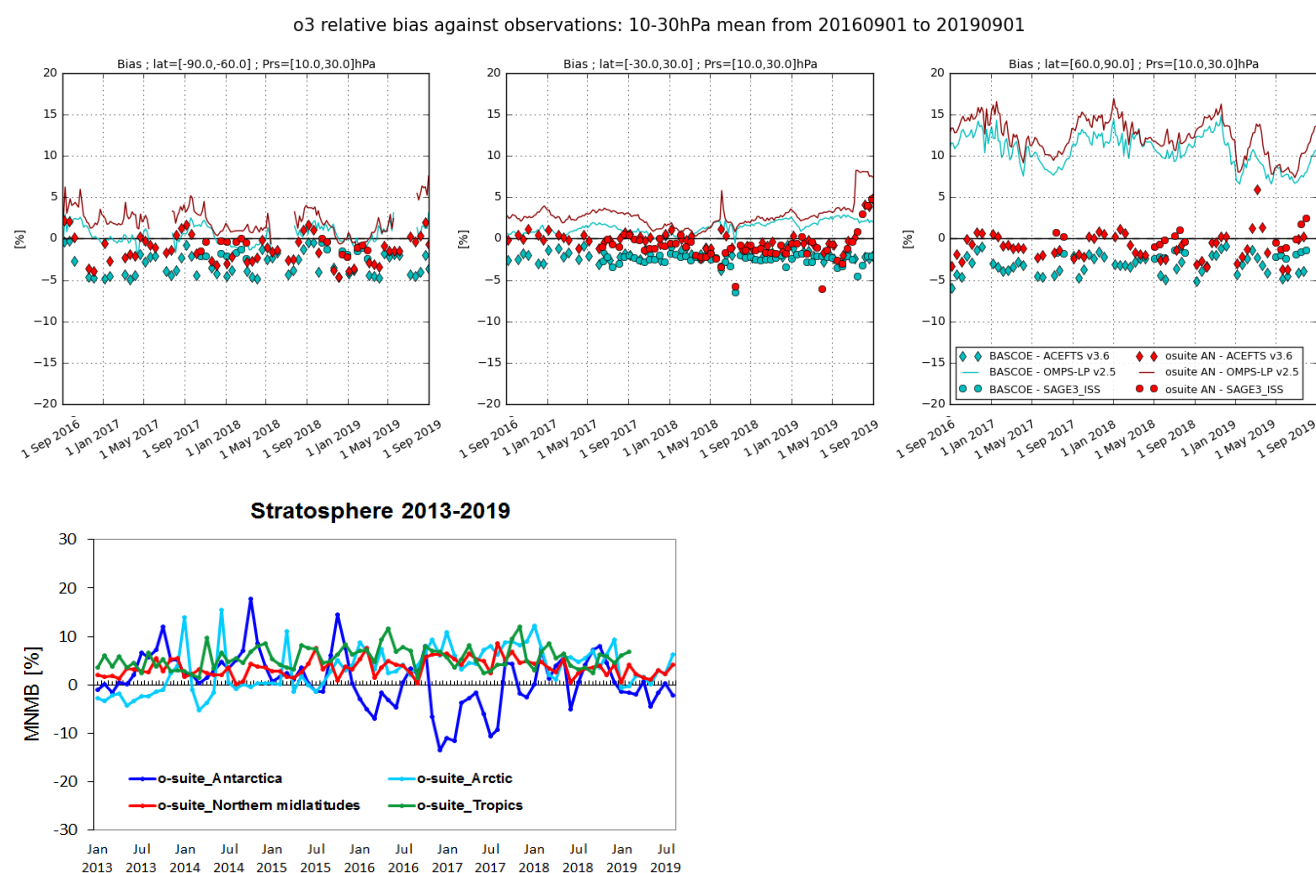o3 relative bias against observations: 10-30hPa mean from 20160901 to 20190901

Figure S.8: Top: Time series comparing models to observations for the period 2016-09-01 to 2019-09-01 in the upper stratosphere (10-30hPa averages): o-suite analyses (red) and BASCOE (cyan) vs OMPS-LP (solid), ACE-FTS (diamonds) and SAGE-III (bullets). Shown is the normalized mean bias (model-obs)/obs (%). Bottom: MNMBs (%) of ozone in the stratosphere from the o-suite against aggregated ozonesonde data in the Arctic (light blue), Antarctic (dark blue) northern midlatitudes (red) and tropics (green) from 2013 to August 2019. The stratosphere is defined as the altitude region between 60 and 10 hPa in the tropics and between 90 and 10 hPa elsewhere.

Ozone layer and UV

Ozone partial columns and vertical profiles

Ozone columns and profiles have been compared with the following observations: vertical profiles from balloon-borne ozonesondes; ground-based remote-sensing observations from the NDACC (Network for the Detection of Atmospheric Composition Change, http://www.ndacc.org); and satellite observations by 3 instruments (OMPS-LP, ACE-FTS and SAGE-III). Furthermore, the o-suite analyses are compared with those delivered by the independent assimilation system BASCOE.

Compared to ozonesondes (Fig. S.8) the model O_3 partial pressures are slightly overestimated in all latitude bands (MNMB between 4 and +12%) except above the Antarctic.

Comparisons with the NDACC network include 20 stations for FTIR, 16 stations for UVVIS stratospheric columns, microwave profiles for Ny Alesund (78.9°N) and Bern (47°N) and LIDAR

profiles at Hohenpeissenberg (47.8°N) and Observatoire Haute Provence (OHP), France (43°N) and Mauna Loa, Hawaii (19.5°N). The comparisons show a general good agreement with the o-suite, with small performance differences between AN and 1d forecasts. At the tropical sites the 1d forecast performs significantly worse since the June 2016 update of the o-suite. This is confirmed by FTIR and the Mauna Loa LIDAR measurements.

The comparison with independent satellite observations (Fig. S.8) is generally in good agreement for the considered period: for ACE-FTS, the NMB is mainly within 10% between 5km and 40km, and mostly within 5% between 15km and 35km except in the tropics; for SAGE-III, the NMB is mainly within 10% between 15km and 40km. OMPS-LP has less regular profiles, but the NMB still remain within 15% for most parts of the 20-40 km range.

Other stratospheric trace gases

Due to the lack of stratospheric chemistry in the C-IFS-CB05 scheme, the only useful product in the stratosphere is ozone. Other species, like NO_2, have also been evaluated but the results are only indicative.

Events

Dust hits Europe during June and July 2019: In summer 2019, MODIS satellite detected two dust outbreaks that reached Central Europe. One outbreak occurred on 22-25 June, originating from Algeria and transported towards the Iberian Peninsula and hitting France, and moving to Eastern Mediterranean the next days. A second dust event was observed on 23-25 July, which also originated in Algeria and Iberian Peninsula, Western France and hitting United Kingdom. Both events were nicely tracked by observations. CAMS AOD o-suite did timely reproduce the spatial distribution of the two dust plumes as shown the comparison with MODIS/Aqua AOD comparison despite the model tends to overestimate the observed maximum values particularly the second event. This second dust outbreak was associated to a deep depression. During the second event, o-suite is predicting PM10 values over 100µg/m^3 in large extensions of the Iberian Peninsula and achieving values up to 50µg/m^3 in United Kingdom.

Smoke over the Arctic from a Siberian fire event: A large fire event, centred at approximately 110°E and 60°N in Siberia north of the Baikal Lake, started on 6 August with large CO emissions for about one week. There was a northwards transport of CO, which reached the Arctic Ocean on 9 August. Elevated concentrations were subsequently found over several parts of the Arctic. The episode is generally well captured by the o-suite, although the total column CO is underestimated compared to Sentinel-5P and IASI satellite observations. The control simulation is more underestimating the CO levels, but the transport episode can still be seen. The episode is not monitored by any surface measurements in Greenland and Svalbard, but a peak in total column CO measured with NDACC FTIR at Thule on 16-17 August may be associated with this transport event.

Table of Contents

1. Introduction

The Copernicus Atmosphere Monitoring Service (CAMS, http://atmosphere.copernicus.eu/) is a component of the European Earth Observation programme Copernicus. The CAMS global near-real time (NRT) service provides daily analyses and forecasts of trace gas and aerosol concentrations. The CAMS near-real time services consist of daily analysis and forecasts with the ECMWF IFS system with data assimilation of trace gas concentrations and aerosol properties. This document presents the system evolution and the validation statistics of the CAMS NRT global atmospheric composition analyses and forecasts. The validation methodology and measurement datasets are discussed in Eskes et al. (2015).

In this report the performance of the system is assessed in two ways: both the longer-term mean performance (seasonality) as well as its ability to capture recent events are documented. Table 1.1 provides an overview of the trace gas species and aerosol aspects discussed in this CAMS near-real time validation report. This document is updated every 3 months to report the recent status of the near-real time service. The report covers results for a period of at least one year to document the seasonality of the biases. Sometimes reference is made to other model versions or the reanalysis to highlight aspects of the near-real time products.

This validation report is accompanied by the "Observations characterization and validation methods" report, Eskes et al. (2018a), which describes the observations used in the comparisons, and the validation methodology. This report can also be found on the global validation page, http://atmosphere.copernicus.eu/user-support/validation/verification-global-services.

Key CAMS NRT products and their users are: Boundary conditions for regional air quality models (e.g. AQMEII, air quality models not participating in CAMS); Long range transport of air pollution (e.g. LRTAP); Stratospheric ozone column and UV (e.g. WMO, DWD); 3D ozone fields (e.g. SPARC). As outlined in the MACC-II Atmospheric Service Validation Protocol (2013) and MACC O-INT document (2011), relevant user requirements are quick looks of validation scores, and quality flags and uncertainty information along with the actual data. This is further stimulated by QA4EO (Quality Assurance Framework for Earth Observation, http://www.qa4eo.org) who write that "all earth observation data and derived products is associated with it a documented and fully traceable quality indicator (QI)". It is our long-term aim to provide such background information. The user is seen as the driver for any specific quality requirements and should assess if any supplied information, as characterised by its associated QI, are "fit for purpose" (QA4EO task team, 2010).

CAMS data are made available to users as data products (grib or netcdf files) and graphical products from ECMWF, accessible through the catalogue on http://atmosphere.copernicus.eu/.

A summary of the system and its recent changes is given in section 2. Subsequent sections give an overview of the performance of the system for various species, and during recent events. Routine validation results can be found online via regularly updated verification pages, http://atmosphere.copernicus.eu/user-support/validation/verification-global-services.
Table 1.2 lists all specific validation websites that can also be found through this link.

Table 1.1: Overview of the trace gas species and aerosol aspects discussed in this CAMS near-real time validation report. Shown are the datasets assimilated in the CAMS analysis (second column) and the datasets used for validation, as shown in this report (third column). Green colours indicate that substantial data is available to either constrain the species in the analysis, or substantial data is available to assess the quality of the analysis. Yellow boxes indicate that measurements are available, but that the impact on the analysis is not very strong or indirect (second column), or that only certain aspects are validated (third column).

Species, vertical range	Assimilation	Validation
Aerosol, optical properties	MODIS Aqua/Terra AOD PMAp AOD	AOD, Ångström: AERONET, GAW, Skynet, MISR, OMI, lidar, ceilometer
Aerosol mass (PM10, PM2.5)	MODIS Aqua/Terra	European AirBase stations
O_3, stratosphere	MLS, GOME-2A, GOME-2B, OMI, SBUV-2, OMPS, TROPOMI	Sonde, lidar, MWR, FTIR, OMPS, ACE-FTS, OSIRIS, BASCOE and MSR analyses
O_3, UT/LS	MLS	IAGOS, ozone sonde
O_3, free troposphere	Indirectly constrained by limb and nadir sounders	IAGOS, ozone sonde, IASI
O_3, PBL / surface		Surface ozone: WMO/GAW, NOAA/ESRL-GMD, AIRBASE
CO, UT/LS	IASI, MOPITT	IAGOS
CO, free troposphere	IASI, MOPITT	IAGOS, MOPITT, IASI, TCCON
CO, PBL / surface	IASI, MOPITT	Surface CO: WMO/GAW, NOAA/ESRL
NO_2, troposphere	OMI, GOME-2, partially constrained due to short lifetime	SCIAMACHY, GOME-2, MAX-DOAS
HCHO		GOME-2, MAX-DOAS
SO_2	GOME-2A, GOME-2B (Volcanic eruptions)	
Stratosphere, other than O_3		NO_2 column only: SCIAMACHY, GOME-2
CO_2, surface, PBL		ICOS
CO_2, column	GOSAT	TCCON
CH_4, surface, PBL		ICOS
CH_4, column	GOSAT, IASI	TCCON

Table 1.2: Overview of quick-look validation websites of the CAMS system.

Reactive gases – Troposphere

IAGOS tropospheric ozone and carbon monoxide:
    http://www.iagos.fr/cams/
 Surface ozone from EMEP (Europe) and NOAA-ESRL (USA):
    http://www.academyofathens.gr/cams
 Tropospheric nitrogen dioxide and formaldehyde columns against satellite retrievals:
    http://www.doas-bremen.de/macc/macc_veri_iup_home.html
 Tropospheric CO columns against satellite retrievals:
    http://www.mpimet-cams.de
 GAW surface ozone and carbon monoxide:
 https://atmosphere.copernicus.eu/charts/cams_gaw_ver/v0d_gaw_oper_operfc_nrt_sites?facets=undefined&time=2018060100,0,2018060100&fieldpair=CO&site=cmn644n00

Reactive gases - Stratosphere

Stratospheric composition:
    http://www.copernicus-stratosphere.eu
 NDACC evaluation in stratosphere and troposphere (the NORS server)
    http://nors-server.aeronomie.be

Aerosol

Evaluation against Aeronet stations:
    http://aerocom.met.no/cams-aerocom-evaluation/
   More in-depth evaluations are available from the Aerocom website.
 WMO Sand and Dust Storm Warning Advisory and Assessment System (SDS-WAS)
   model intercomparison and evaluation:
   http://sds-was.aemet.es/forecast-products/models
 Aeronet verification of CAMS NRT forecasts:
   https://atmosphere.copernicus.eu/charts/cams_aeronet_ver/?facets=undefined&time=2019020100,0,2019020100&site=ARM_Graciosa

Satellite data monitoring
Monitoring of satellite data usage in the Near-Real-Time production: https://atmosphere.copernicus.eu/charts/cams/cams_satmon?facets=undefined&time=2016071800&Parameter=AURA_MLS_profile_Ozone_1_GLOBE

Naming and color-coding conventions in this report follow the scheme as given in Table 1.3.

Table 1.3. Naming and colour conventions as adopted in this report.

Name in figs	experiment	Colour
{obs name}	{obs}	black
o-suite D+0 FC	0001	red
control	gsyg	blue
GHG high-resolution run	gqpe / ghqy	orange
GHG global analysis	gqiq	green

2. System summary and model background information

The specifics of the different CAMS model versions are given below (section 2.1) including an overview of model changes. Other systems used in CAMS are listed in section 2.2. An overview of products derived from this system is given in section 2.3. Timeliness and availability of the CAMS products is given in section 2.4.

2.1 System based on the ECMWF IFS model (the o-suite and control run)

Key model information is given on the CAMS data-assimilation and forecast run o-suite and its control experiment, used to assess the performance of the assimilation. The forecast products are listed in Table 2.1. Table 2.2 provides information on the satellite data used in the o-suite. Further details on the different model runs and their data usage can be found at
http://atmosphere.copernicus.eu/documentation-global-systems.

Table 2.1: Overview of model runs assessed in this validation report.

Forecast system	Exp. ID	Brief description	Upgrades (e-suite ID)	Cycle
o-suite	0001	Operational CAMS DA/FC run	20190709-present 20180626-20190708 20170926-20180625 20170124-20170926 20160621-20170124 20150903-20160620 20140918-20150902	46R1 45R1 43R3 43R1 41R1 41R1 40R2
Control	h7c4 gzhy gsyg gnhb gjjh geuh g4o2	control FC run without DA	20190709-present 20180626-20190708 20170926-20180625 20170124-20170926 20160621-20170124 20150901-20160620 20140701-20150902	46R1 45R1 43R3 43R1 41R1 41R1 40R2
GHG run	ghqy gf39	High resolution T1279, NRT CO_2 and CH_4 without DA	20160301-20170621 20150101-20160229	
	gqpe	High resolution Tco1279 (~9km) NRT CO_2, CH_4 and linCO forecast, initialized from GHG analysis gqiq and CAMS operational CO analysis	20170101-present	
	gqiq	GHG analysis Tco399 (~25km)	20170101-present	

Table 2.2: Satellite retrievals of reactive gases and aerosol optical depth that are actively assimilated in the o-suite.

Instrument	Satellite	Provider	Version	Type	Status
MLS	AURA	NASA	V4	O3 Profiles	20130107 -
OMI	AURA	NASA	V883	O3 Total column	20090901 -
GOME-2A	Metop-A	Eumetsat	GDP 4.8	O3 Total column	20131007 - 20181231
GOME-2B	Metop-B	Eumetsat	GDP 4.8	O3 Total column	20140512 -
SBUV-2	NOAA-19	NOAA	V8	O3 21 layer profiles	20121007 -
OMPS	Suomi-NPP	NOAA / EUMETSAT		O3 Profiles	20170124 - 20190409
TROPOMI	Sentinel-5P	ESA		O3 column	20181204-
IASI	MetOp-A	LATMOS/ULB Eumetsat	-	CO Total column	20090901 - 20180621 20180622 -
IASI	MetOp-B	LATMOS/ULB Eumetsat	-	CO Total column	20140918 - 20180621 20180622 -
MOPITT	TERRA	NCAR	V5-TIR V7-TIR V7-TIR Lance V8-TIR	CO Total column	20130129 - 20160124 - 20180626 20180626 20190702
OMI	AURA	KNMI	DOMINO V2.0	NO2 Tropospheric column	20120705 -
GOME-2A/2B	METOP A/B	Eumetsat	GDP 4.8	NO2 Tropospheric column	20180626 -
OMI	AURA	NASA	v003	SO2 Tropospheric column	20120705-20150901
GOME-2A/2B	METOP A/B	Eumetsat	GDP 4.8	SO2 Tropospheric column	20150902 -
MODIS	AQUA / TERRA	NASA	Col. 5 Deep Blue Col. 6, 6.1	Aerosol total optical depth, fire radiative power	20090901 - 20150902 - 20170124 -
PMAp	METOP-A METOP-B	EUMETSAT		AOD	20170124 - 20170926 -

2.1.1 The CAMS o-suite

The o-suite consists of the IFS-CB05 chemistry combined with the CAMS bulk aerosol model. The chemistry is described in Flemming et al. (2015) and Flemming et al. (2017), aerosol is described in Morcrette et al. (2009). The forecast length is 120 h. The o-suite data is stored under **expver '0001'** of **class 'MC'**. On 21 June 2016 the model resolution has seen an upgrade from T255 to T511, and forecasts are produced twice per day.
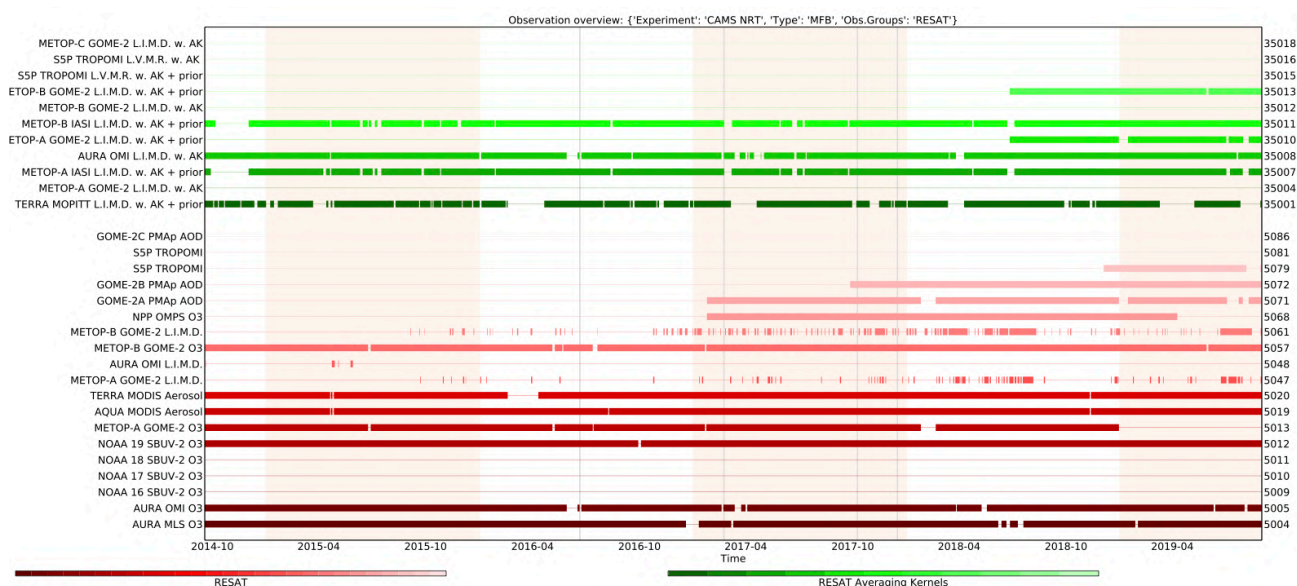
Figure 2.1: Satellite observation usage in the real-time analysis, for ozone, CO, aerosol AOD and NO_2, from October 2014 onwards. Top eight rows: products assimilated using averaging kernels. New assimilated products since the 24 January 2017 upgrade are the PMaP AOD including GOME-2B and OMPS ozone profile observations. Sentinel-5P TROPOMI ozone is assimilated since Dec. 2018 ($5079=O_3$) and other products are monitored ($35016=NO_2$, $35015=CO$, $5081=SO_2$).

A short summary of the main model specifications:

- The modified CB05 tropospheric chemistry is used (Williams et al., 2013), originally taken from the TM5 chemistry transport model (Huijnen et al., 2010)
- Stratospheric ozone during the forecast is computed from the Cariolle scheme (Cariolle and Teyssèdre, 2007) as already available in IFS, while stratospheric NOx is constrained through a climatological ratio of HNO_3/O_3 at 10 hPa.
- Monthly mean dry deposition velocities are based on the SUMO model provided by the MOCAGE team.
- Data assimilation is described in Inness et al. (2015) and Benedetti et al. (2009) for chemical trace gases and aerosol, respectively. Satellite data assimilated is listed in Table 2.2 and Fig. 2.1.
- Anthropogenic and biogenic emissions are based on MACCity (Granier et al., 2011) and a climatology of the MEGAN-MACC emission inventories (Sindelarova et al., 2014)
- NRT fire emissions are taken from GFASv1.2 (Kaiser et al. 2012).

The aerosol model includes 14 prognostic variables (Remy et al., 2019).

- 3 size bins each for sea-salt and desert dust

- 2 bins (hydrophibic and hydrophobic) each for organic matter and black carbon

- 1 bin for sulphate

- 2 bins (fine and coarse) for nitrate (New since 46R1)

- 1 bin for ammonium (New since 46R1)

The SO$_2$ precursor for sulphate aerosol no longer exists as a separate prognostic in the aerosol scheme, which since 46R1 couples directly to the SO$_2$ in the chemistry scheme instead. Likewise, the precursors for the new nitrate and ammonium aerosol (nitric acid and ammonia) are also part of the chemistry scheme rather than the aerosol scheme.

Aerosol total mass is constrained by the assimilation of MODIS AOD (Benedetti et al. 2009). A variational bias correction for the MODIS AOD is in place based on the approach used also elsewhere in the IFS (Dee and Uppala, 2009).

A history of updates of the o-suite is given in Table 2.3, and is documented in earlier MACC-VAL and CAMS reports: https://atmosphere.copernicus.eu/node/326. This includes a list with changes concerning the assimilation system.

The CAMS o-suite system is upgraded regularly, following updates to the ECMWF meteorological model as well as CAMS-specific updates such as changes in chemical data assimilation. These changes are documented in e-suite validation reports, as can be found from the link above. Essential model upgrades are also documented in Table 2.3.

The penultimate upgrade of the system (45r1) took place on 26 June 2018. This upgrade is also relevant for this report (for the period up to 8 July), and the validation for this upgrade is described in Eskes et al., 2018b/2018c.

2.1.2 Short description of the latest CAMS upgrade (46r1)

The latest upgrade of the system took place on 9 July 2019 and is based on IFS version cy46r1_CAMS and involves the move from 60 to 137 vertical levels, see https://atmosphere.copernicus.eu/cycle-46r1 or https://confluence.ecmwf.int/display/COPSRV/Current+global+production+suites.

The validation for this 46r1 upgrade is described in Basart et al., 2019: https://atmosphere.copernicus.eu/sites/default/files/2019-07/CAMS84_2018SC1_D3.2.1-201907_esuite_v1.pdf

The meteorological changes can be found on the ECMWF-IFS CY46R1 page, https://confluence.ecmwf.int/display/FCST/Implementation+of+IFS+cycle+46R1.

The atmospheric composition content of the new cycle includes the following aspects:

Assimilation:

- New model-error covariance matrices for aerosol and chemistry at 137 levels.

Observations:

- No new atmospheric composition observations compared to Cycle 45r1.

Emissions:

- New emissions inventories: CAMS_GLOB_ANT v2.1 (anthropogenic) and CAMS_GLOB_BIO v1.1 (biogenic), in place of previous MACCity and MEGAN_MACC inventories.

Table 2.3: Long-term o-suite system updates.

Date	o-suite update
2009.08.01	Start of first NRT experiment f7kn with coupled MOZART chemistry, without aerosol. Also without data assimilation.
2009.09.01	Start of first MACC NRT experiment f93i, based on meteo cy36r1, MOZART v3.0 chemistry, MACC aerosol model, RETRO/REAS and GFEDv2 climatological emissions, T159L60 (IFS) and 1.875°×1.875° (MOZART) resolution.
2012.07.05	Update to experiment fnyp: based on meteo cy37r3, MOZART v3.5 chemistry, where changes mostly affect the stratosphere, MACCity (gas-phase), GFASv1 emissions (gas phase and aerosol), T255L60 (IFS) and 1.125°×1.125° (MOZART) resolution. Rebalancing aerosol model, affecting dust.
2013.10.07	Update of experiment fnyp from e-suite experiment fwu0: based on meteo cy38r2, no changes to chemistry, but significant rebalancing aerosol model. Assimilation of 21 layer SBUV/2 ozone product
2014.02.24	Update of experiment fnyp from e-suite experiment fzpr: based on meteo cy40r1. No significant changes to chemistry and aerosol models.
2014.09.18	Update to experiment g4e2: based on meteo cy40r2. In this model version IFS-CB05 is introduced to model atmospheric chemistry.
2015.09.03	Update to experiment g9rr: based on meteo cy41r1.
2016.06.21	Update to experiment 0067: based on meteo cy41r1, but a resolution increase from T255 to T511, and two production runs per day
2017.01.24	Update to cycle 43R1_CAMS, T511L60
2017.09.26	Update to cycle 43R3_CAMS, T511L60
2018.06.26	Update to cycle 45R1_CAMS, T511L60
2019-07-09	Update to cycle 46R1_CAMS, T511L137

- Biomass-burning injection heights from GFAS and updated diurnal cycle. In particular, this reduces the overestimation of near-surface PM2.5 during fire events.

- Anthropogenic SOA production was updated with a diurnal cycle and a regionally-varying ratio to CO emissions. This has a small impact on AOD, but significantly reduces night-time near-surface PM2.5 in polluted regions.

- New online dust emission scheme (Nabat et al., 2012). This increases total dust emissions and shifts them towards larger particle sizes, in line with recent literature. An updated dust source function improves the selection of source regions, reducing "gaps" in dust emissions.

- Sea-salt production over freshwater lakes eliminated. This corrects an issue that was particularly noticeable over the Great Lakes.

Other model changes:

- Vertical resolution increased from 60 levels to 137 levels, matching that used at ECMWF for NWP. This includes moving the model top from 0.1 hPa to 0.01 hPa.

- New nitrate and ammonium aerosol species are included and are coupled to the gas-phase nitrogen chemistry. This is a major expansion of the aerosol species represented in the model, giving a more complete representation of the species which contribute to e.g. PM2.5 over Europe.

- Sulphur species (SO_2 and SO_4) coupled between chemistry and aerosol schemes. See discontinued parameters below. This brings a greater consistency between the chemistry and aerosol products related to the sulphur cycle.

- Online calculation of dry deposition velocities for trace gases. This was already in place for aerosols in 45r1 and allows the deposition scheme to better account for variations in surface properties.

- Updates to wet deposition parameterisations. This brings improvements in the distinction between scavenging by liquid and ice and harmonises the treatment for aerosols and trace gases.

- Updates to chemical reaction rates following latest recommendations by JPL/IUPAC.

2.1.3 Control

The control run (relevant expver = **gzhy,** since 26/06/2018; expver = **h7c4** since 09/07/2019) applies the same settings as the respective o-suites, based on the coupled IFS-CB05 system with CAMS aerosol for cy54r1, except that data assimilation is not switched on. The only exception with regard to this setup are:

- The full meteorology in the control run is also initialized from the ECMWF operational NWP analyses. Note that this is different from the o-suite, which uses its own data assimilation setup for meteorology. This can cause slight differences in meteorological fields between o-suite and control, e.g. as seen in evaluations of upper stratospheric temperatures.

2.1.4 High-resolution CO_2 and CH_4 forecasts and delayed-mode analyses

The pre-operational forecasts of CO_2 and CH_4 use an independent setup of the IFS at a resolution of TL1279, i.e. ~16 km horizontal, and with 137 levels. This system runs in real time and does not apply data assimilation for the greenhouse gases.

The land vegetation fluxes for CO_2 are modelled on-line by the CTESSEL carbon module (Boussetta et al., 2013). A biogenic flux adjustment scheme is used in order to reduce large-scale biases in the net ecosystem fluxes (Agusti-Panareda, 2015). The anthropogenic fluxes are based on the annual mean EDGARv4.2 inventory using the most recent year available (i.e. 2008) with estimated and climatological trends to extrapolate to the current year. The fire fluxes are from GFAS (Kaiser et al., 2012). Methane fluxes are prescribed in the IFS using inventory and climatological data sets, consistent with those used as prior information in the CH_4 flux inversions from Bergamaschi et al. (2009). The anthropogenic fluxes are from the EDGAR 4.2 database (Janssens-Maenhout et al, 2012)

valid for the year 2008. The biomass burning emissions are from GFAS v1.2 (Kaiser et al., 2012). The high-resolution forecast experiments also included a linear CO scheme (Massart et al., 2015).

The experiments analysed in this report are:

- "**ghqy**" from March 2016. The initial conditions used in ghqy on 1st of March 2016 are from the GHG analysis (experiment gg5m). Furthermore, the meteorological analysis used to initialize the ghqy forecast changed resolution and model grid in March 2016. Note that the CO_2, CH_4 and linear CO tracers are free-running.

- "**gqpe**" from January 2017 to present. It runs with a TCO1279 Gaussian cubic octahedral grid (equivalent to approximately 9km horizontal resolution). Note that the CO_2, CH_4 and linear CO tracers are initialized with the GHG analysis (gqiq) for CO_2 and CH_4 and the CAMS operational analysis for CO.

- The greenhouse gas analysis experiment "**gqiq**" runs on a TCO399 grid (equivalent to around 25km) and 137 vertical levels and is available from January 2017. This experiment runs in delayed mode (4 days behind real time) and makes use of observations from TANSO-GOSAT (methane and CO_2) and MetOp-IASI (methane).

2.2 Other systems

2.2.1 BASCOE

The NRT analyses and forecasts of ozone and related species for the stratosphere, as delivered by the Belgian Assimilation System for Chemical ObsErvations (BASCOE) of BIRA-IASB (Lefever et al., 2014; Errera et al., 2008), are used as an independent model evaluation of the CAMS products. The NRT BASCOE product is the ozone analysis of Aura/MLS-SCI level 2 standard products, run in the following configuration (version 05.07):

- The following species are assimilated: O_3, H_2O, HNO_3, HCl, HOCl, N_2O and ClO.
- It lags by typically 4 days, due to latency time of 4 days for arrival of non-ozone data from Aura/MLS-SCI (i.e. the scientific offline Aura/MLS dataset).
- Global horizontal grid with a 3.75° longitude by 2.5° latitude resolution.
- Vertical grid is hybrid-pressure and consists in 86 levels extending from 0.01 hPa to the surface.
- Winds, temperature and surface pressure are interpolated in the ECMWF operational 6-hourly analyses.
- Time steps of 20 minutes, output every 3 hours

See the stratospheric ozone service at http://www.copernicus-stratosphere.eu/. It delivers graphical products dedicated to stratospheric composition and allows easy comparison between the results of o-suite, BASCOE and TM3DAM. The BASCOE data products (HDF4 files) are also distributed from this webpage. Other details and bibliographic references on BASCOE can be found at http://bascoe.oma.be/. A detailed change log for BASCOE can be found at http://www.copernicus-stratosphere.eu/4_NRT_products/3_Models_changelogs/BASCOE.php.

2.2.2 TM3DAM and the multi-sensor reanalysis

One of the MACC products was a 30-year reanalysis, near-real time analysis and 10-day forecast of ozone column amounts performed with the KNMI TM3DAM data assimilation system, the Multi-Sensor Reanalysis (MSR) system (van der A et al., 2010, 2013),
http://www.temis.nl/macc/index.php?link=o3_msr_intro.html.
The corresponding validation report can be found at
 http://www.copernicus-atmosphere.eu/services/gac/global_verification/validation_reports/.

The NRT TM3DAM product used for the validation of the CAMS NRT streams is the ozone analysis of Envisat/SCIAMACHY (until April 2012), AURA/OMI, and MetOp-A/GOME-2, run in the following configuration:

- total O_3 columns are assimilated
- Global horizontal grid with a 3° longitude by 2° latitude resolution.
- Vertical grid is hybrid-pressure and consists in 44 levels extending from 0.1 hPa to 100 hPa.
- Dynamical fields from ECMWF operational 6-hourly analysis.

An update of the MSR (MSR-2) was presented in van der A et al. (2015), which extended the record to 43 years based on ERA-interim reanalysis meteo and with an improved resolution of 1x1 degree.

2.2.3 SDS-WAS multimodel ensemble

The World Meteorological Organization's Sand and Dust Storm Warning Advisory and Assessment System (WMO SDS-WAS) for Northern Africa, Middle East and Europe (NAMEE) Regional Center (http://sds-was.aemet.es/) has established a protocol to routinely exchange products from dust forecast models as the basis for both near-real-time and delayed common model evaluation. Currently, twelve regional and global models (see the complete list in the following link https://sds-was.aemet.es/forecast-products/forecast-evaluation/model-inter-comparison-and-forecast-evaluation/at_download/file) provides daily operational dust forecasts (i.e. dust optical depth, DOD, and dust surface concentration).

Different multi-model products are generated from the different prediction models. Two products describing centrality (multi-model median and mean) and two products describing spread (standard deviation and range of variation) are daily computed. In order to generate them, the model outputs are bi-linearly interpolated to a common grid mesh of 0.5º x 0.5º. The multimodel DOD (at 550 nm) Median from nine dust prediction models participating in the SDS-WAS Regional Center is used for the validation of the CAMS NRT streams.

2.3 CAMS products

An extended list of output products from the NRT stream o-suite are available as 3-hourly instantaneous values up to five forecast days. These are available from ECMWF (through ftp in grib2 and netcdf format, https://atmosphere.copernicus.eu/data).

2.4 Availability and timing of CAMS products

The availability statistics provided in Table 2.6 are computed for the end of the 5-day forecast run.

Table 2.6: Timeliness of the o-suite from December 2014. From June 2016 onwards CAMS has produced two forecasts per day.

Months	On time, 10 & 22 utc	80th perc	90th perc	95th perc
Dec-Feb '14-'15	97%	D+0, 19:43	D+0, 20:28	D+0, 21:13
Mar-May 2015	96%	D+0, 19:38	D+0, 21:03	D+0, 21:40
Jun-Aug 2015	95%	D+0, 20:24	D+0, 20:53	D+0, 21:54
Sept-Nov 2015	95%	D+0, 19:44	D+0, 20:55	D+0, 21:51
Dec-Feb '15-'16	100%	D+0, 18:39	D+0, 18:57	D+0, 19:43
Mar-May 2016	98%	D+0, 19:32	D+0, 19:47	D+0, 20:00
Jun-Aug 2016 (00 and 12 cycle)	100%	D+0, 08:53 D+0, 20:55	D+0, 09:04 D+0, 21:01	D+0, 09:18 D+0, 21:18
Sep-Nov 2016	98.9%	D+0, 08:44 D+0, 20:44	D+0, 08:51 D+0, 20:48	D+0, 08:52 D+0, 20:51
Dec 2016 - Feb 2017	99.4%	D+0, 09:02 D+0, 21:01	D+0, 09:11 D+0, 21:02	D+0, 09:18 D+0, 21:04
Mar-May 2017	100%	D+0, 09:08 D+0, 21:07	D+0, 09:14 D+0, 21:09	D+0, 09:19 D+0, 21:11
Jun-Aug 2017	100%	D+0, 09:05 D+0, 21:05	D+0, 09:07 D+0, 21:08	D+0, 9:09 D+0, 21:10
Sep-Nov 2017	100%	D+0, 09:02 D+0, 21:00	D+0, 09:05 D+0, 21:04	D+0, 9:09 D+0, 21:07
Dec 2017 - Feb 2018	98.33%	D+0, 08:55 D+0, 20:54	D+0, 08:59 D+0, 20:59	D+0, 09:01 D+0, 21:02
Mar-May 2018	98.9%	D+0, 09:00 D+0, 21:00	D+0, 09:06 D+0, 21:03	D+0, 09:08 D+0, 21:06
Jun-Aug 2018	100%	D+0, 09:11 D+0, 21:07	D+0, 09:14 D+0, 21:09	D+0, 09:20 D+0, 21:11
Sep-Nov 2018	100%	D+0, 09:05 D+0, 21:03	D+0, 09:09 D+0, 21:07	D+0, 09:13 D+0, 21:10
Dec 2018 - Feb 2019	98.85%	D+0, 09:03 D+0, 21:04	D+0, 09:06 D+0, 21:06	D+0, 09:08 D+0, 21:10
Mar-May 2019	100%	D+0, 09:07 D+0, 21:05	D+0, 09:10 D+0, 21:09	D+0, 09:12 D+0, 21:11
Jun-Aug 2019	99.46%	D+0, 09:19 D+0, 21:14	D+0, 09:22 D+0, 21:17	D+0, 09:27 D+0, 21:19

The CAMS production KPI is defined as the percentage of cycles in which all the general data dissemination tasks are completed before the deadlines: 10 UTC for the 00:00 and 22 UTC for the 12:00 UTC run. This was in part based on requirements from the regional models. We note that at present most regional models can still provide their forecasts even if the global forecast is available a bit later. Note that since 21 June 2016 two CAMS forecasts are produced each day.

The o-suite data delivery for the period June-August 2019 (JJA-2019) was very good, with an on-time percentage of 99.46 %. Only on 21 August there was a 7 min delay with some of the products.

3. Tropospheric Ozone

3.1 Validation with sonde data in the free troposphere

Model profiles of the CAMS runs were compared to free tropospheric balloon sonde measurement data of 38 stations taken from the NDACC, WOUDC, NILU and SHADOZ databases for January 2013 to August 2019 (see Fig. 3.1.1 - 3.1.2). Towards the end of the period, the number of available soundings decreases, which implies that the evaluation results may become less representative. The figures contain the number of profiles in each month that are available for the evaluation. The methodology for model comparison against the observations is described in Eskes et al., 2018a. The free troposphere is defined as the altitude range between 750 and 200hPa in the tropics and between 750 and 300hPa elsewhere.

MNMBs for the o-suite are mostly within the range ±20%, for all months, in all zonal bands, except for the Tropics and Antarctica, where larger positive MNMBs up to ±45% appear, see Fig. 3.1.4. During the last year (August 2018 to August 2019) MNMBs are ±18% over the Arctic and Northern Midlatitudes and up to 30% for Antarctica and the Tropics, see Fig. 3.1.1.-3.1.4.
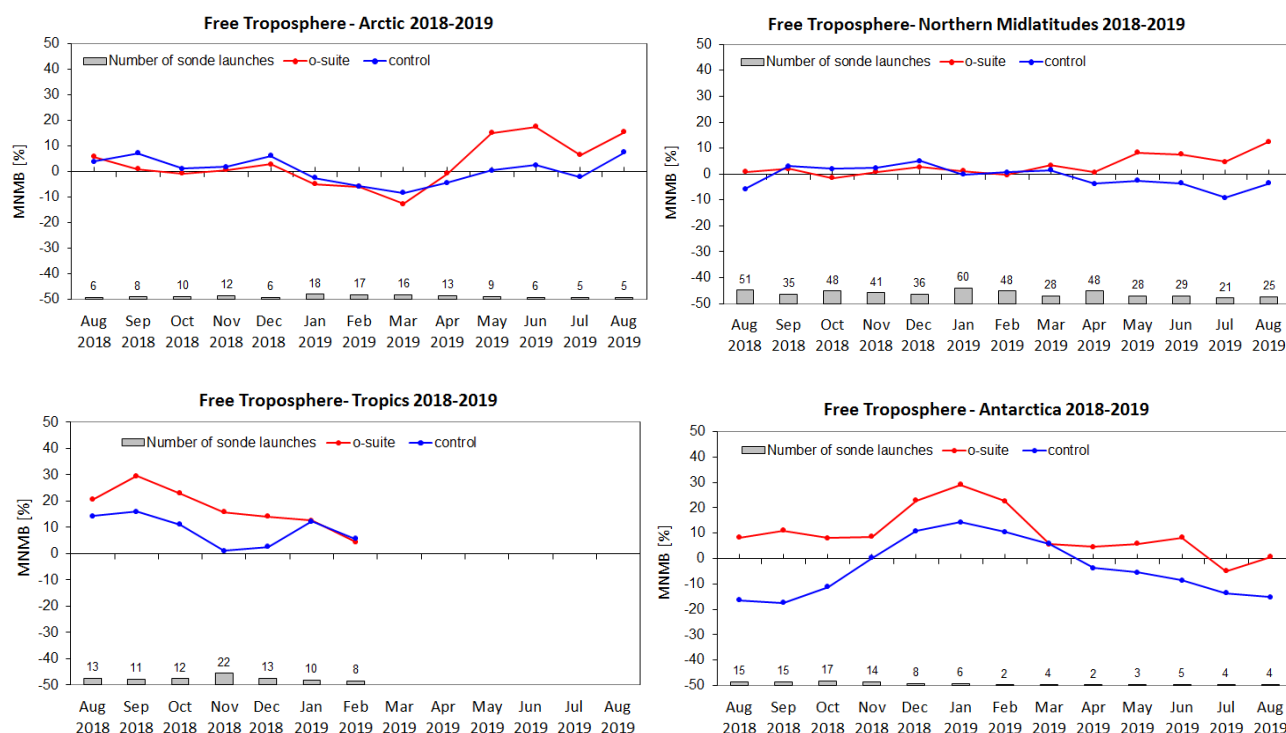
Figure 3.1.2: MNMBs (%) of ozone in the free troposphere (between 750 and 200hPa (Tropics) / 300hPa) from the IFS model runs against aggregated sonde data over the Arctic (top-left) and the Northern mid latitudes (top-right), Tropics (bottom-left) and Antarctica (bottom-right). The numbers indicate the amount of individual number of sondes.
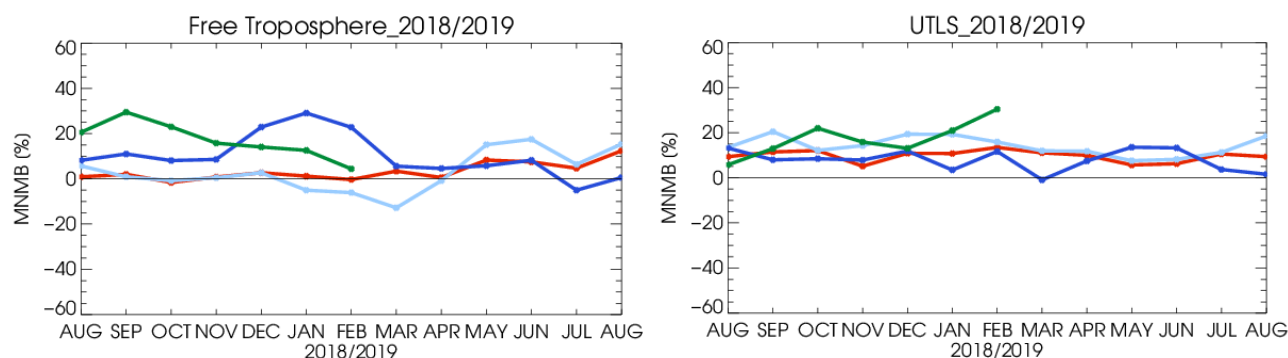
Figure 3.1.3: MNMBs (%) of ozone in the free troposphere (left, between 750 and 200hPa (Tropics) / 300hPa) and UTLS (right, between 300 and 100hPa / 300-60hPa tropics) from the IFS model runs against aggregated sonde data over the Tropics (green) and Antarctica (blue), Arctic (light blue) and Northern Midlatitudes (red).
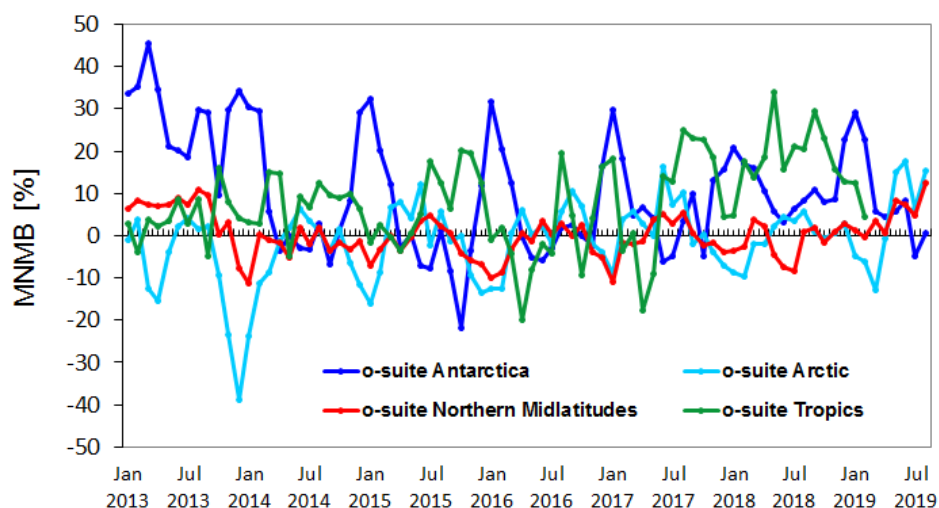
Figure 3.1.4: Time series of MNMB of free-tropospheric ozone (layer between 750 and 300 hP) in the o-suite, compared against ozonesondes, averaged over different latitude bands.

Over the Arctic, the o-suite mostly shows slightly positive MNMBs during summer and spring (MNMBs up to 6%), while during the winter season the MNMBs get negative (within -13%). From May 2019 onwards, however, the stopping of the assimilation of OMPS data leads to larger positive biases up to 18%, see Fig. 3.1.1.

Over the NH mid-latitudes MNMBs for the o-suite are on average close to zero all year round (usually within ±5%), which is generally a clear improvement compared to the control run, which shows larger MNMBs during the respective period. Same as over the Arctic, there is an increase of biases from May 2019 onwards (to up to 12.4%), due to changes in the data assimilation.

Over the Tropics and over Antarctica, ozone mixing ratios are mostly overestimated by the o-suite (up to 30%) by the o-suite, see Fig. 3.1.2. The control run shows large negative MNMBs for Antarctica.

In the UTLS, ozone is overestimated by the o-suite over all regions. MNMBs range mostly within ±20%.
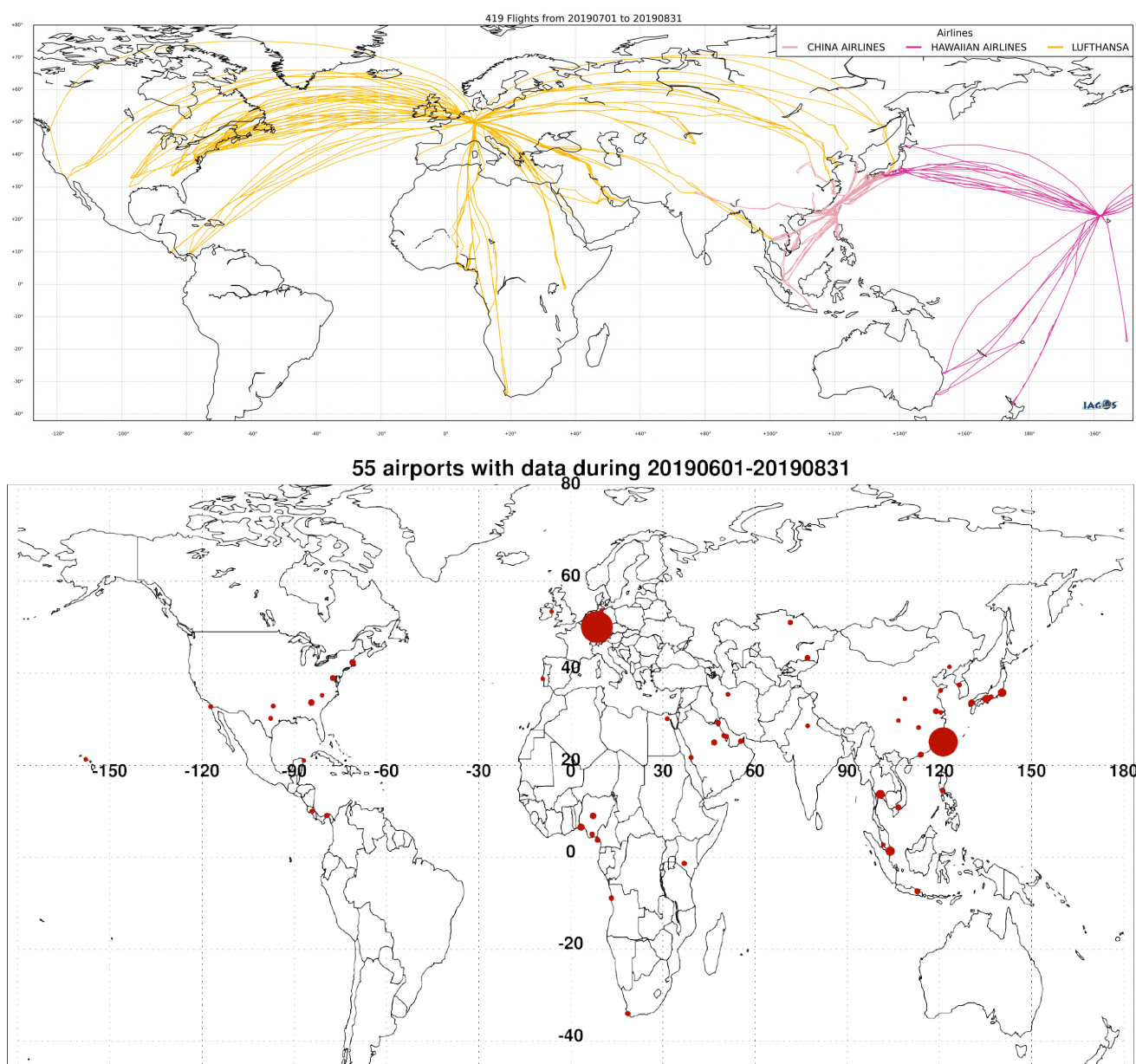
Figure 3.2.1. Map of the flights (top) and the visited airports (bottom) during the period July - August 2019, by the IAGOS-equipped aircraft. The size of the plotting circle represents the number of profiles available.

3.2 Ozone validation with IAGOS data

The daily profiles of ozone measured at airports around the world are shown on the CAMS website at http://www.iagos-data.fr/cams/nrt_profiles.php. For the period from June - August 2019, the measurements displayed on the web pages and in this report include only the data as validated by the instrument PI. The available flights and available airports are shown in Fig. 3.2.1 top and bottom respectively. Performance indicators have been calculated for different parts of the IAGOS operations.
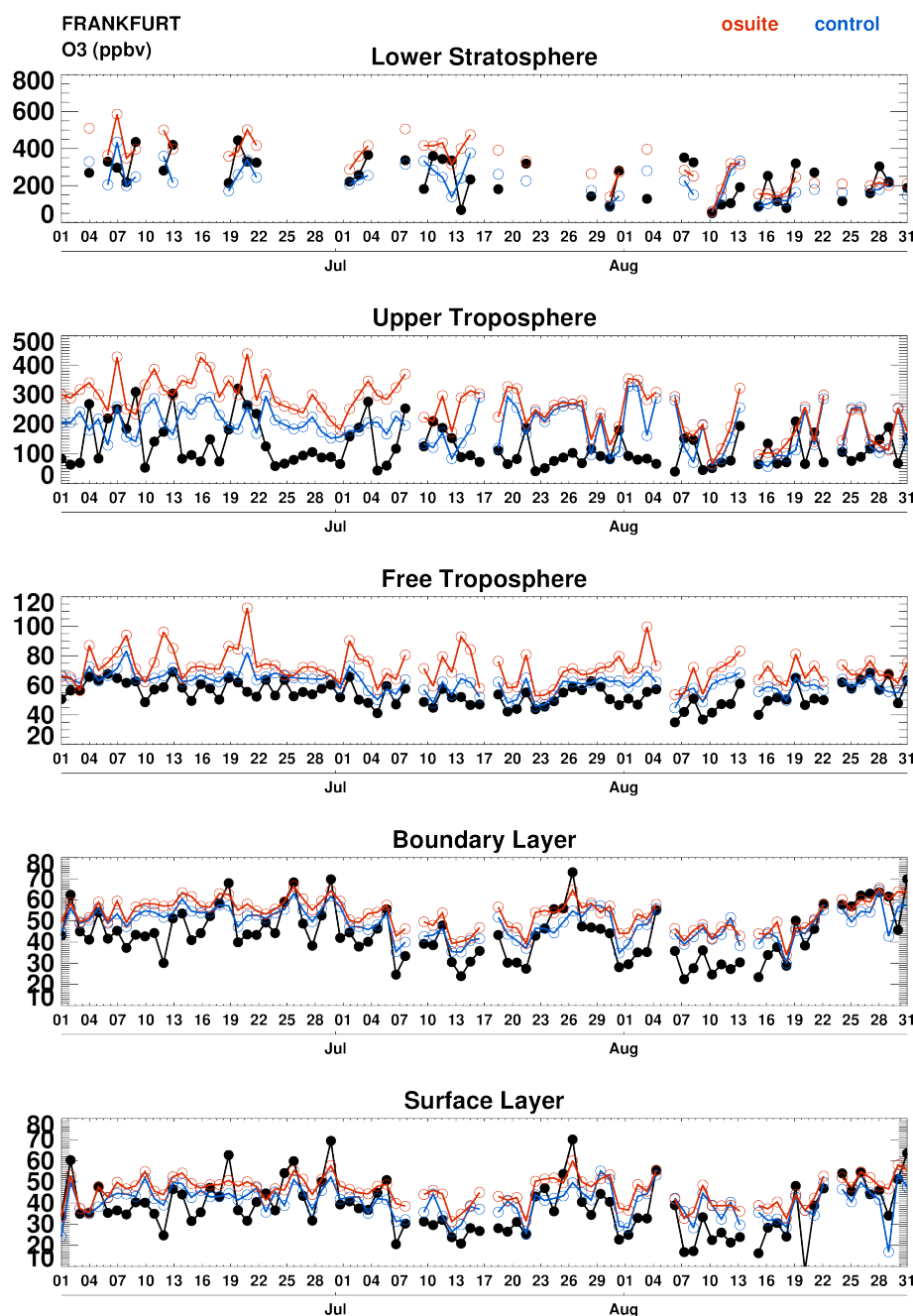
Figure 3.2.2. Time series of daily mean ozone over Frankfurt during JJA 2019 for 5 layers: Surface Layer, Boundary Layer, Free Troposphere, Upper Troposphere and Lower Stratosphere. The o-suite is shown in red and associated control run in blue.

Six aircraft were operating during this period. With these aircrafts, operating fully over the three-month period, we can expect a total of about 1260 flights. The actual number of flights within the period was 419 (838 profiles) giving a performance of 33 %. These flights are shown in Fig. 3.2.1 (top). Forty seven percent (47%) of the operational flights had usable measurements of ozone and 46% of the flights had usable CO. Delivering these O_3 and CO data were two aircraft from Lufthansa operating from Frankfurt, and one from China Airlines. Fig. 3.2.1 (bottom) shows the available airports, with a plotting circle scaled to the highest number of flights at an airport.
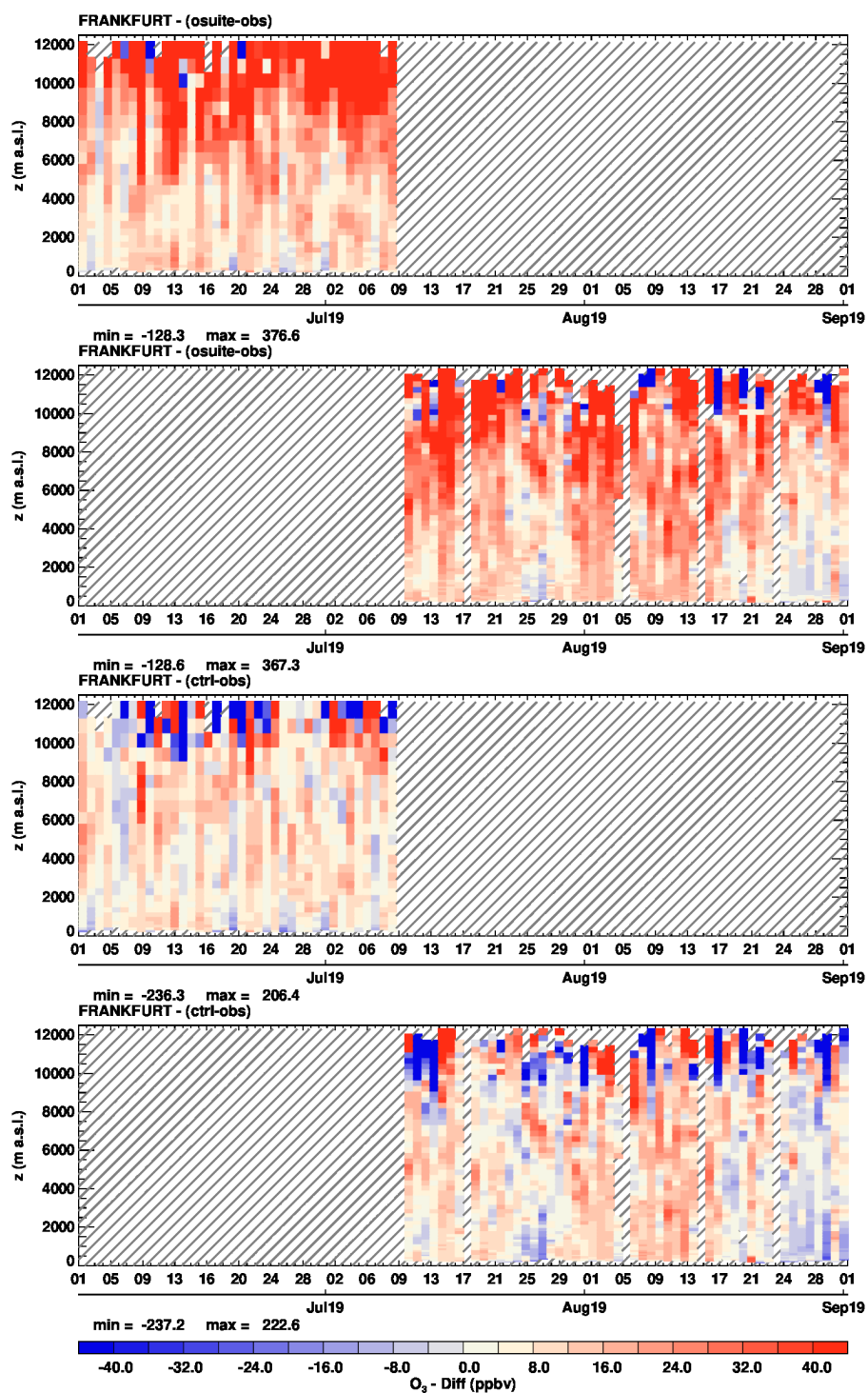
Figure 3.2.3 Time series of the absolute differences (model - observations) in daily profiles for ozone over Frankfurt during JJA 2019. Two upper panels correspond to o-suite (before and after upgrade) and two lower panels to control run (before and after upgrade).

Europe

Fig. 3.2.2 presents ozone time series at Frankfurt during the full period June-August 2019 for 5 atmospheric layers. Time series of the profile differences are also presented in Fig. 3.2.3. Due to the change in the number of levels these time series are split in two sub-periods for the period before and after the upgrade of the CAMS global o-suite to the new 137-level version.

As shown in this figure there is no major difference in the bias between these two periods. However, there are some discontinuities in the bias values which are present especially after the upgrade and for both runs. Indeed, for some profiles large overestimations are found sometimes from the surface up to the free troposphere. This can also be seen on Fig. 3.2.2 though less clearly with on the one hand high ozone episodes in the surface and boundary layer that are generally well represented by the models and low ozone values that are often largely overestimated by the models. These large values of the bias might not be related to the upgrade as they have already been seen and mentioned in the previous MAM 2019 reports. This is also consistent with the results obtained in the JJA 2019 regional report. This bias is attributed to an issue with the ozone sensors on board one of the LH aircraft and improves after 16 August 2019 when the instrument was replaced. In the UTLS region, ozone is mostly overestimated by the o-suite throughout the full period while the performance of control run is better with smaller overestimations in the UT layer and underestimations in the LS layer.

Some examples of individual profiles are presented in Fig. 3.2.4a-b. During the first heat wave in 2019 over Europe which took place in the last week of June, several profiles show high ozone values in the low troposphere at Frankfurt. On 26 and 30 June, ozone mixing ratios between 90 and 100 ppbv have been measured by IAGOS in the surface and boundary layer (Fig. 3.2.4a). As shown by the time series, the two ozone peaks observed in these layers on 26 and 30 June are simulated by the models but underestimated by both runs with a slightly better performance for the o-suite. On 19 June another peak is observed by IAGOS but not clearly reproduced by the model (Fig. 3.2.2). Some profiles are also available during the second heat wave over Europe in late July (see Fig. 3.2.4b). On 26 July, the ozone mixing ratio is constant in the surface and boundary layer with a value of about 90 ppbv. On this date the profile observed at 15:06 shows an ozone enhancement above the boundary layer between 2000 and 3000 m. Like for the episode of June, according to the timeseries the models simulate the episode (Fig. 3.2.2), but the major peak is underestimated by both configurations, again with a better performance from the o-suite.
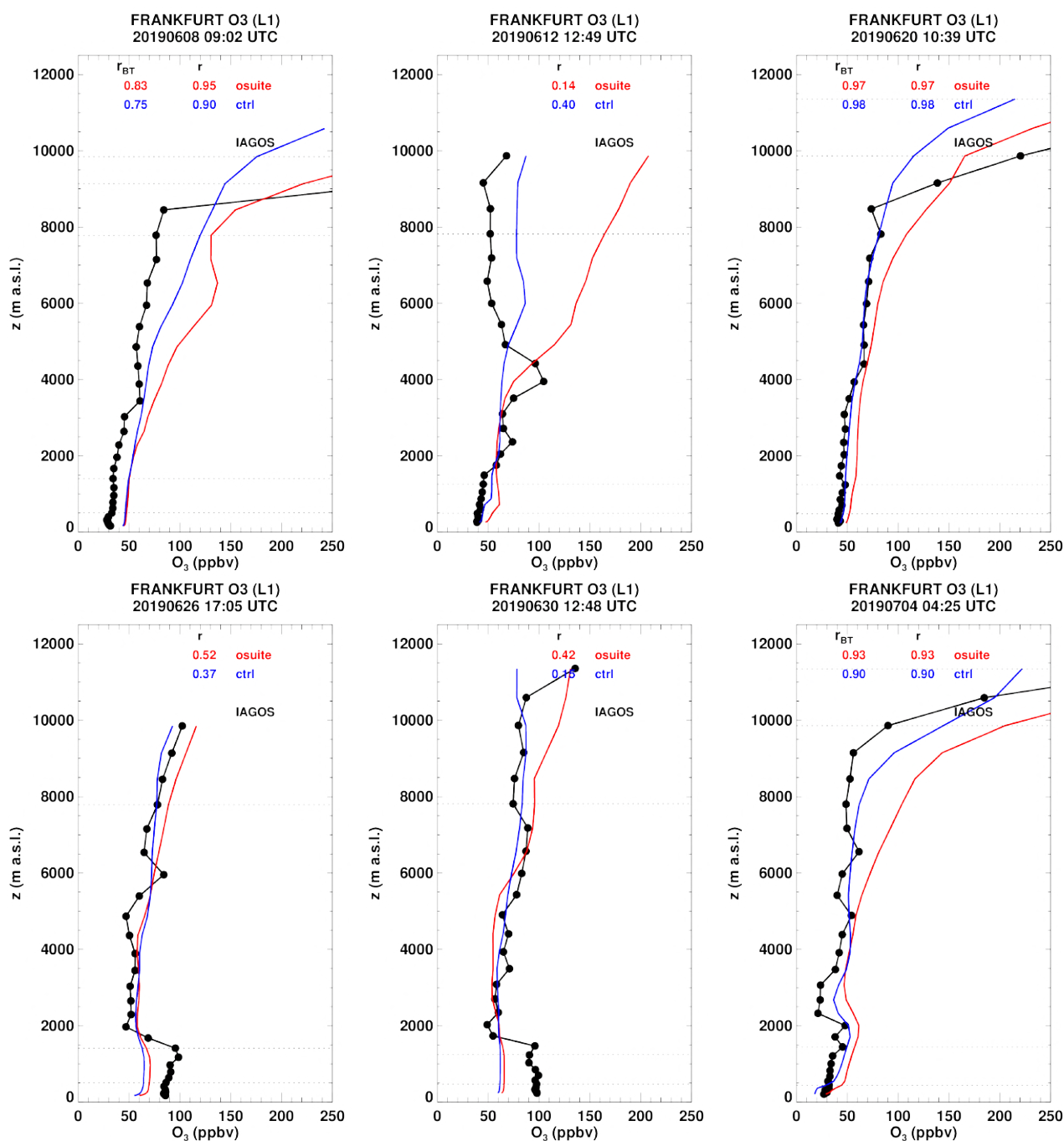
Figure 3.2.4a. Daily profiles for ozone from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over Europe during JJA 2019.
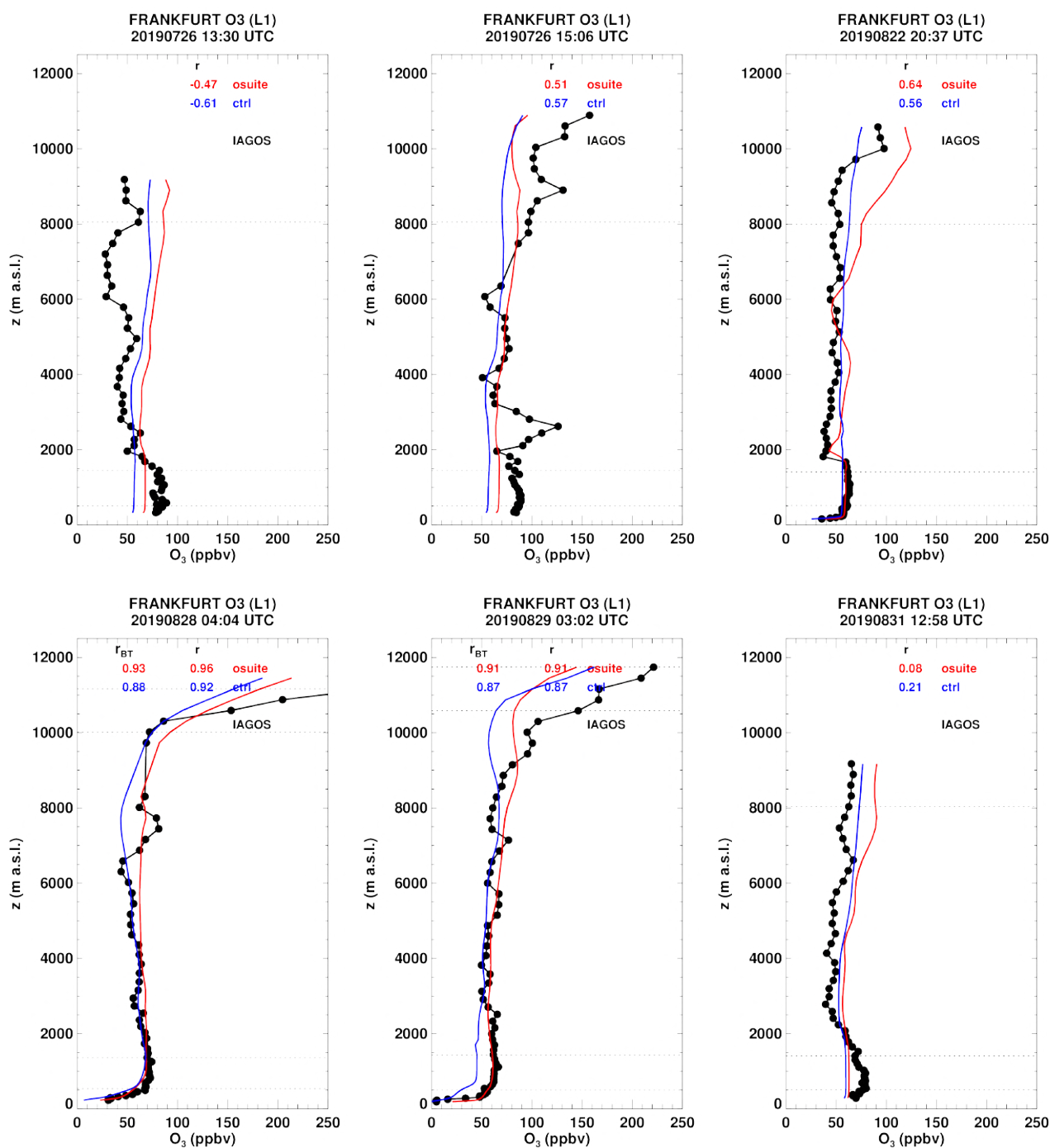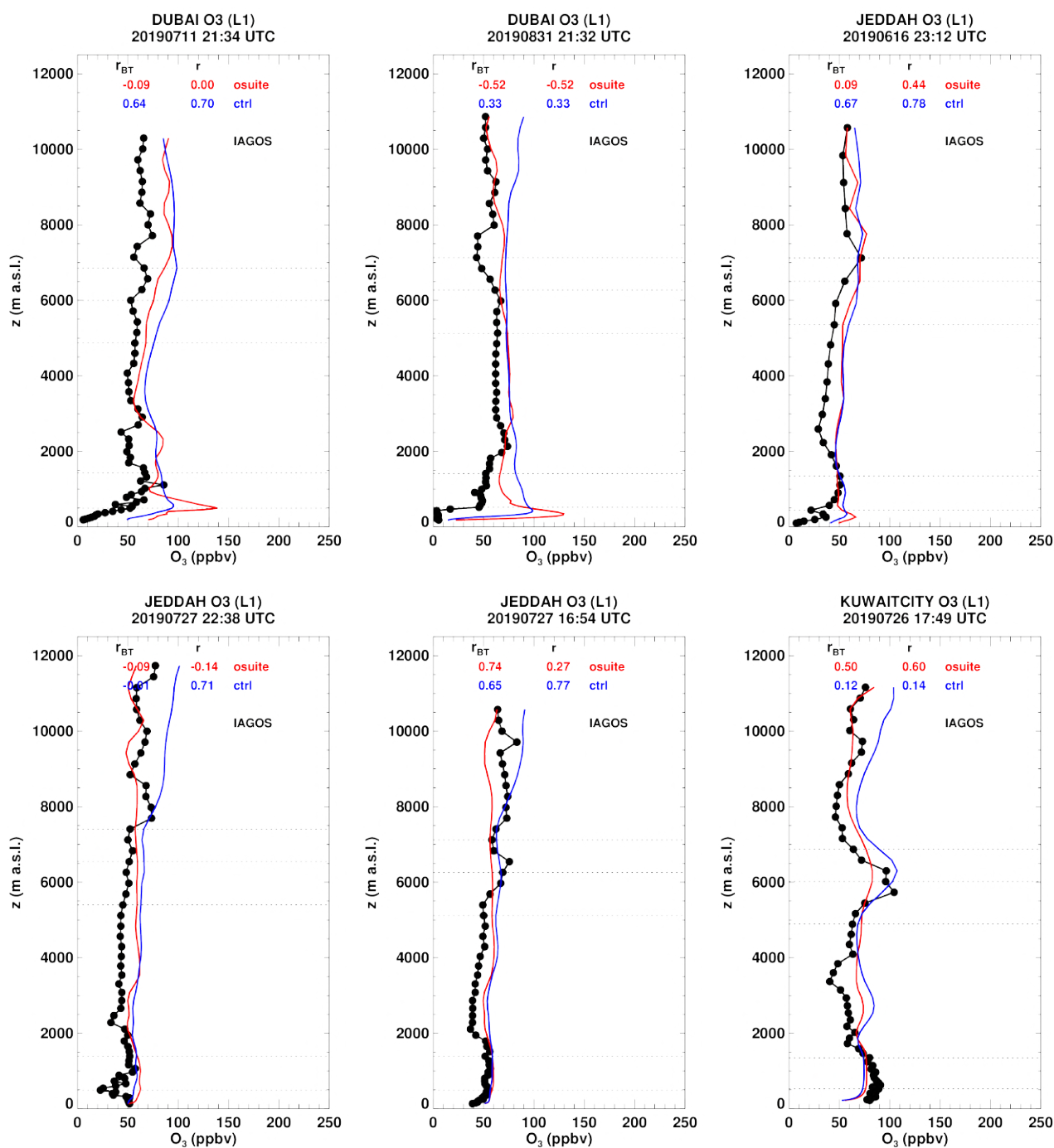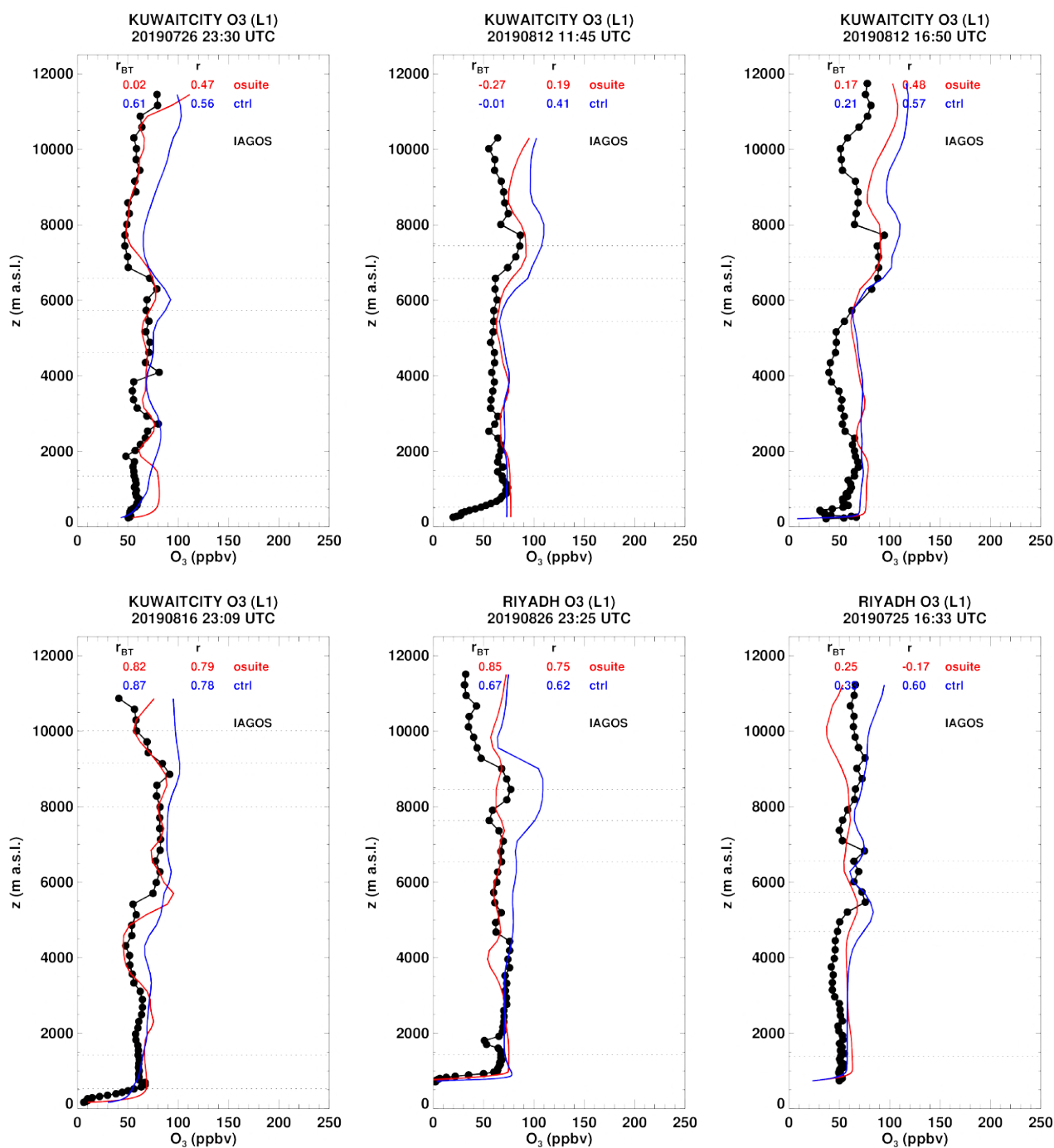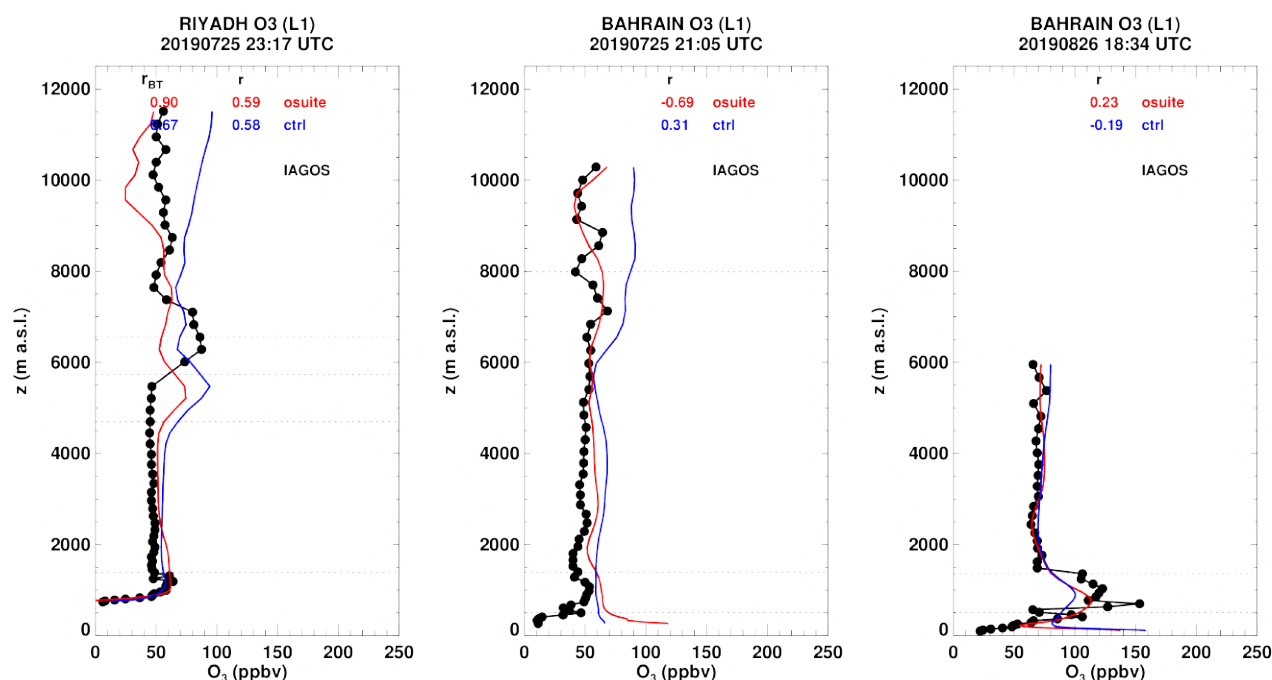
Figure 3.2.4b. Daily profiles for ozone from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over Europe during JJA 2019.

At the end of August (Fig. 3.2.2), high values of ozone are observed in the surface and boundary layer. Profiles for 28, 29 and 31 August are presented in Fig.3.2.4. These values are well reproduced by both runs with a slightly better agreement for the o-suite on day 29.

In the UTLS several peaks of ozone are observed in the time series during the first half of June (Fig. 3.2.2), as well as around 20 June and 4 July. Related profiles are also presented in Fig. 3.2.4a-b. On the profile for 8 June, 20 June and 4 July the tropopause is found between 8000 and 10000 m and this is well reproduced by the models with a better performance from the o-suite in these two profiles of June and conversely a better performance from control run in the profiles of July. On 12 June, a maximum in ozone values if observed in the free troposphere at an altitude of 4000 m. Although the o-suite detects an increase in ozone values in the UTLS this maximum is not reproduced by either of the models.

Middle East

Profiles are available over the Middle East at the airports of Dubai, Jeddah, Kuwait City, Riyadh and Bahrain (Fig. 3.2.5a-b). For these locations, ozone is in general in the range of 50-70 ppbv in the surface and boundary layer with the exception of a profile at Bahrain on 26 August where ozone values present a maximum of 150 ppbv in the boundary layer. In general, the results from both runs are very similar with a rather good agreement in the surface and boundary layer at most airports of the Middle East. At Bahrain the two profiles show that surface values are largely overestimated. Moreover, at the airport of Dubai boundary layer values are largely overestimated (by almost 50%).
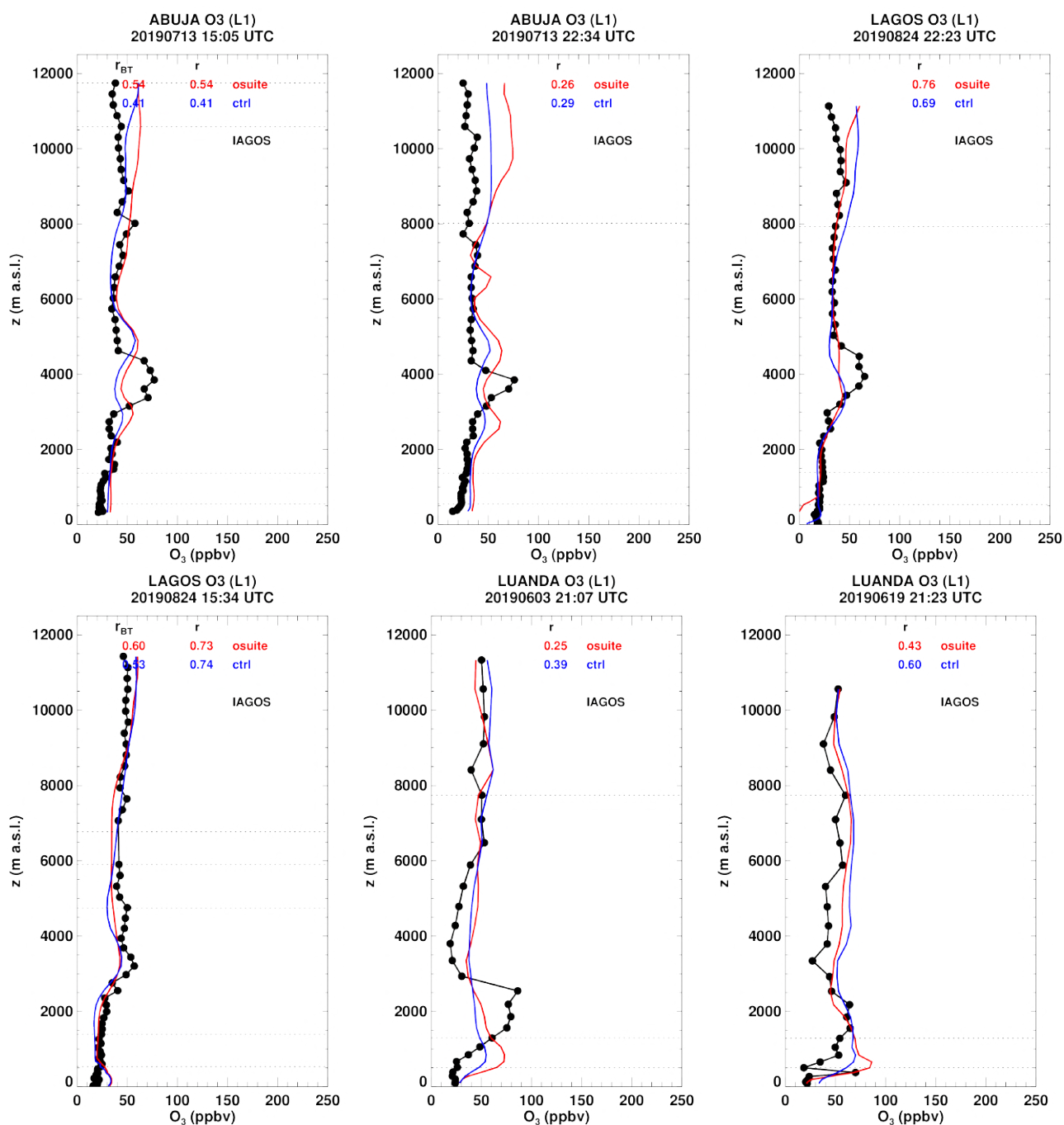
Figure 3.2.5a. Daily profile for ozone from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over the Middle East during JJA 2019.
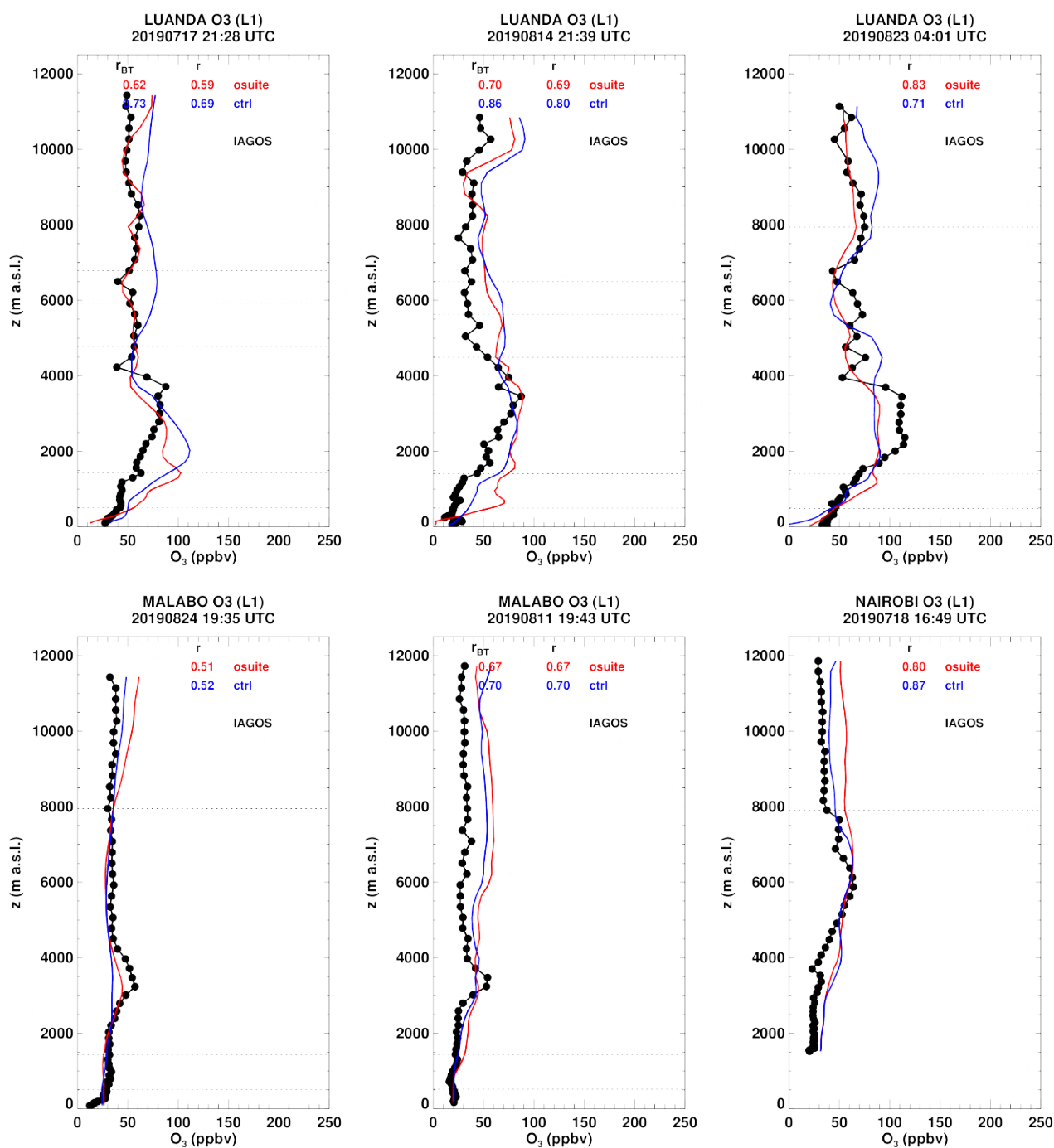
Figure 3.2.5b. Daily profile for ozone from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over the Middle East during JJA 2019.
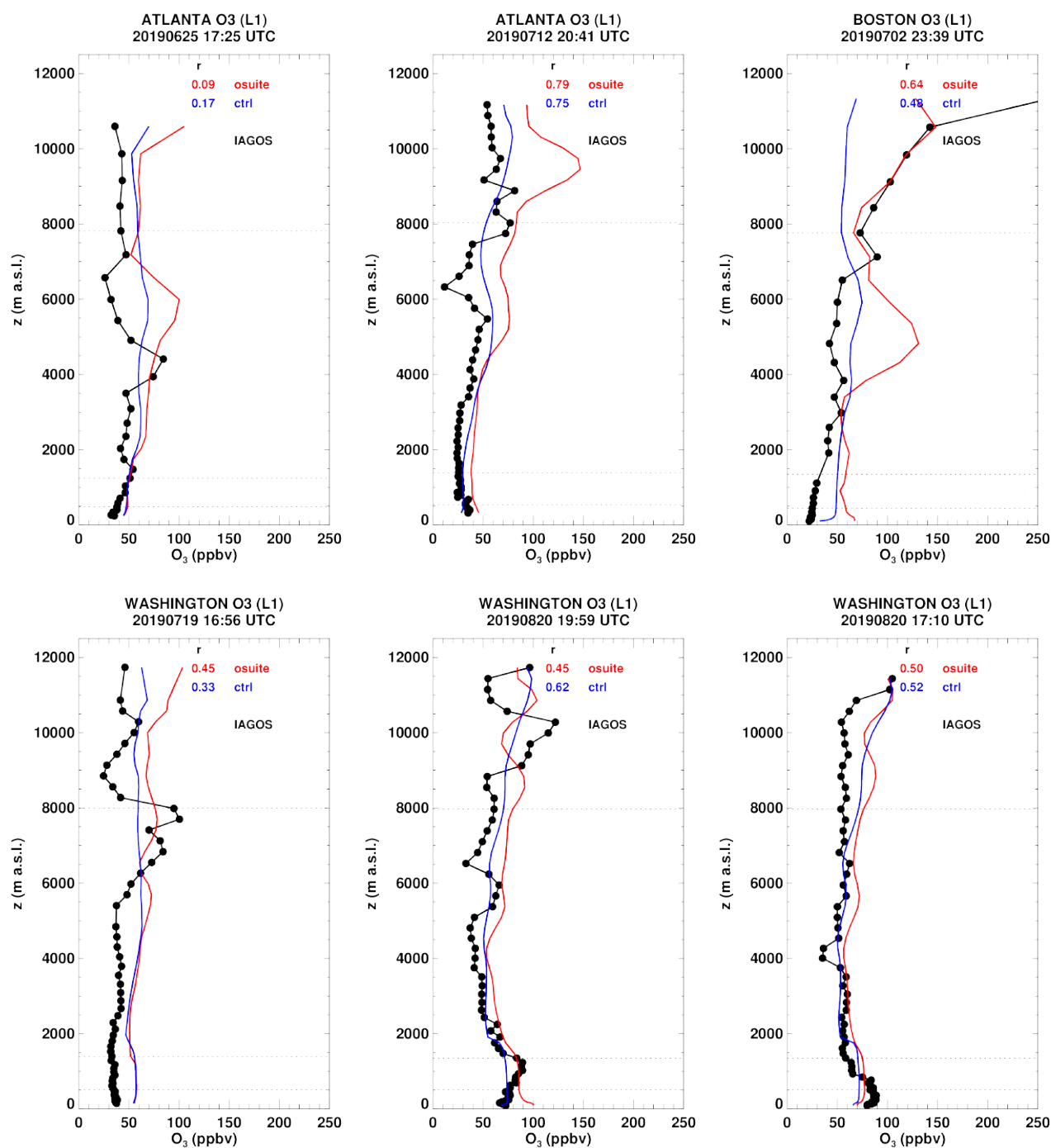
Figure 3.2.5c. Daily profile for ozone from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over the Middle East during JJA 2019.

Africa

IAGOS profiles are available at Nigerian airports of Abuja and Port Harcourt, in the Gulf of Guinea at Lagos and Malabo, as well as at the Kenyan airport of Nairobi (Fig. 3.2.6). At these locations surface and boundary layer concentrations are below 50 ppbv and in general both CAMS configurations agree well with observations. Many profiles present maxima in the free troposphere, see e.g. the measurements at Luanda. These maxima are related to plumes from forest fires as they are correlated to maxima in CO at the same altitude. Both CAMS runs behave rather similarly and reproduce the shape of these profiles, but the altitude of the maxima is often slightly different from those observed, and their magnitude is either overestimated or underestimated.

North America

Many North American airports are visited during JJA 2019: Atlanta, Boston, Washington, Dallas, Austin and San Diego. For all profiles, ozone values are generally near or below 50 ppbv from the surface to the free troposphere (Fig. 3.2.7a-b). The results from both runs are very similar from the surface to the mid-troposphere with a rather good agreement, except in Dallas, Atlanta and Boston where large overestimations are found in the free troposphere. In the free troposphere a maximum is observed over Atlanta on 25 June and over Washington on 19 July. In Atlanta this maximum is simulated by the o-suite but at a higher altitude, whereas in Washington no maximum is found in the models. In Boston, for the profile of 2 July, a maximum of ozone is found near 5000 m which could be related to the transport of pollution from Artic forest fires (see CO section). However, this increase in ozone is not observed by IAGOS and neither in the control run results. In the UTLS, the bias is large for both runs and the behaviour often differs between the two models.

Figure 3.2.6a. Daily profiles for ozone from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over West Africa during JJA 2019.

Figure 3.2.6b. Daily profiles for ozone from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over West Africa during JJA 2019.

Figure 3.2.7.a Daily profiles for ozone from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over North America during JJA 2019.
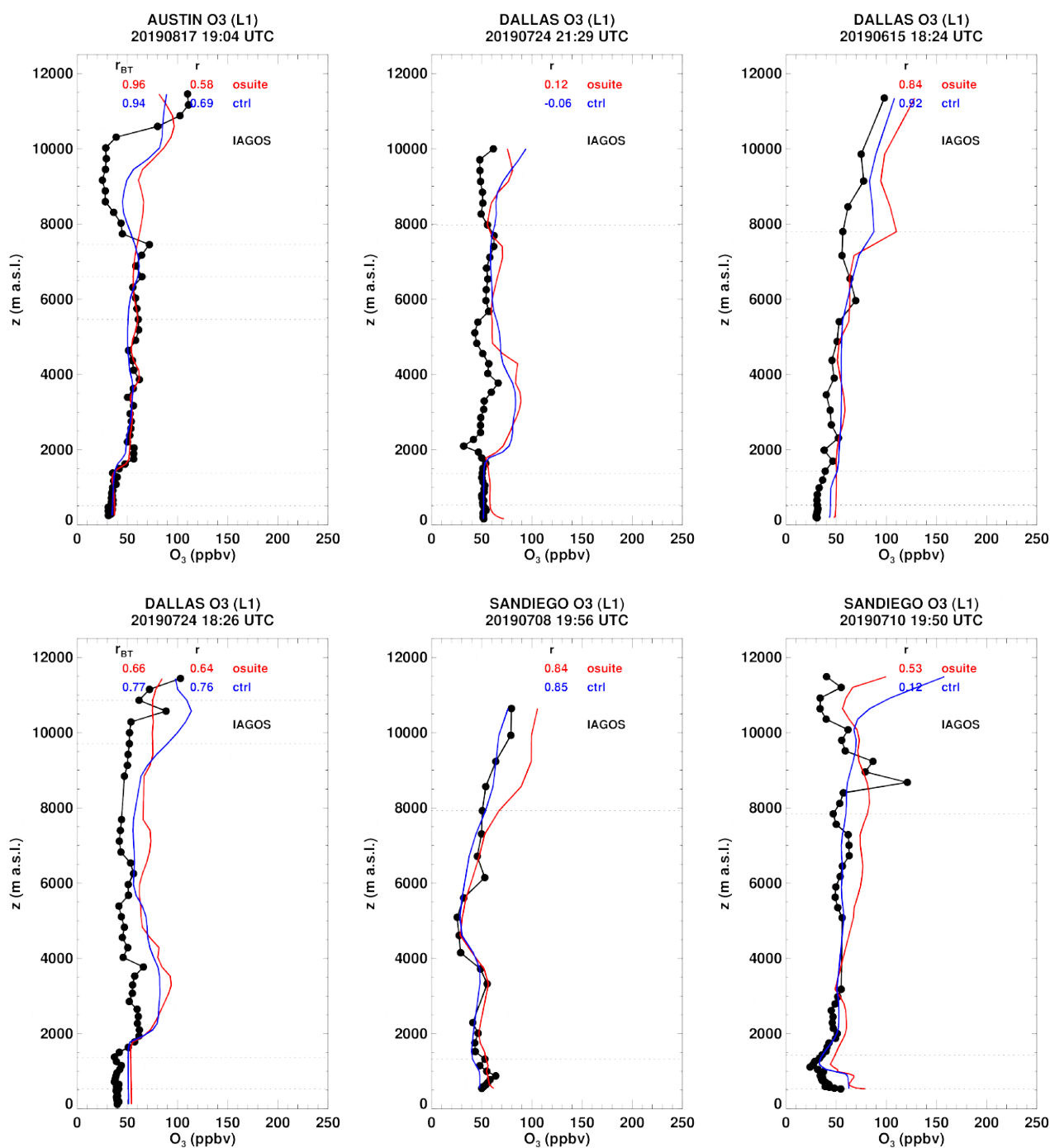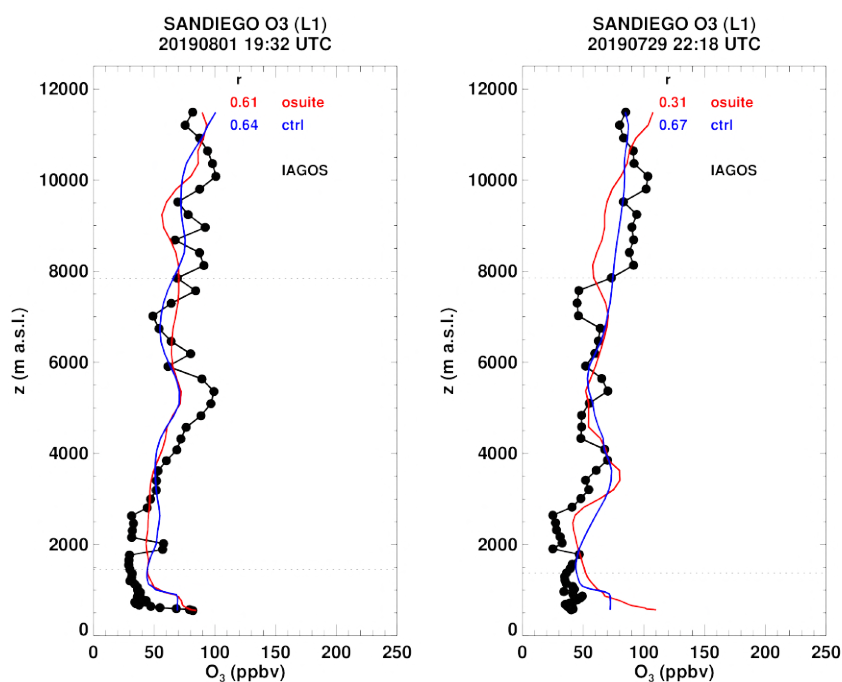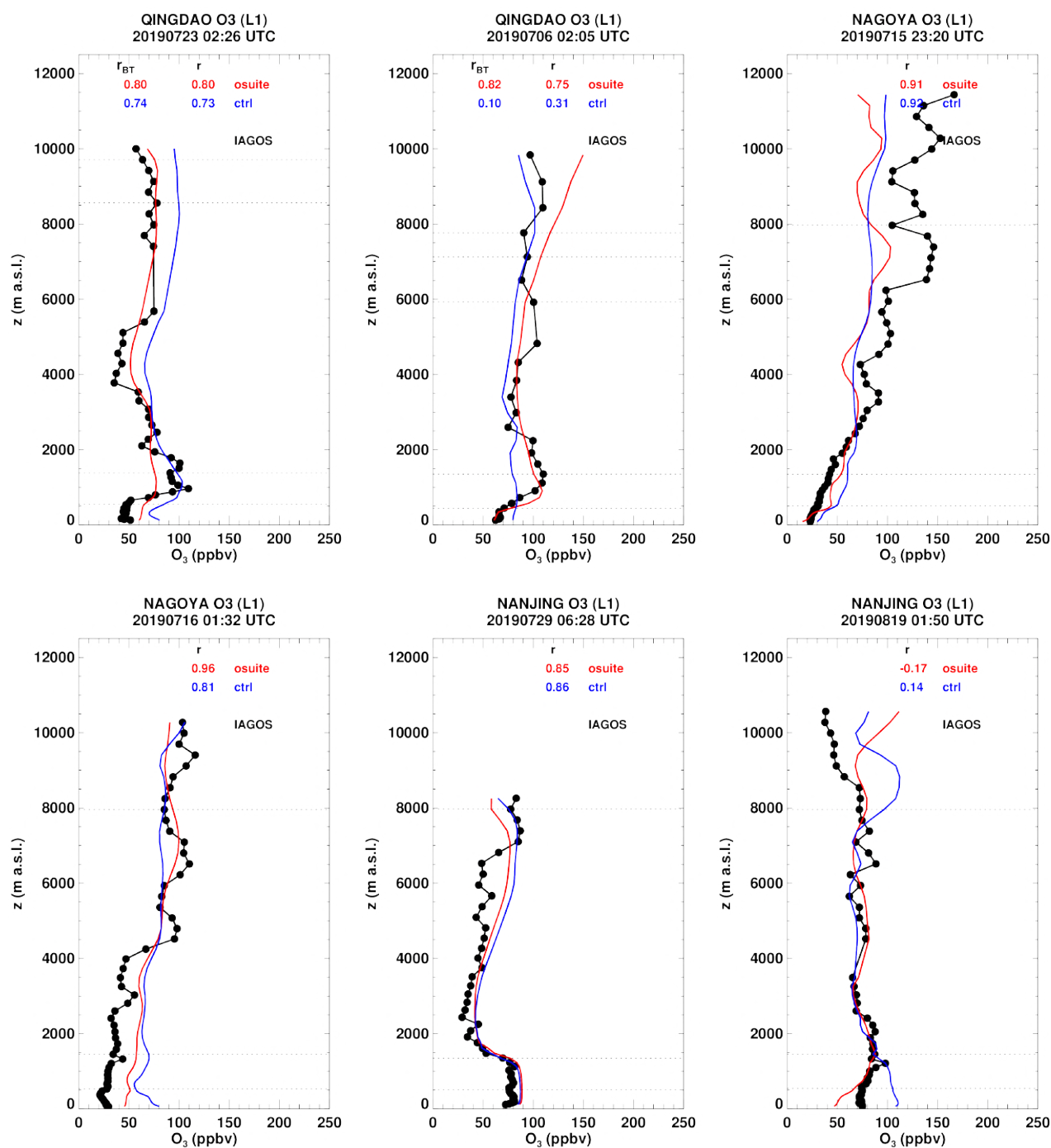
Figure 3.2.7.b Daily profiles for ozone from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over North America during JJA 2019.

Figure 3.2.7.c Daily profiles for ozone from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over North America during JJA 2019.

East Asia

Over East Asia ozone profiles are available at Qingdao, Nagoya and Nanjing. These profiles have very different shapes and range of values for ozone. The results from the o-suite and control run are often very similar in the low to mid-troposphere. In the boundary layer, when ozone values are high (>100 ppbv), sometimes o-suite and sometimes the control shows a better performance. In the UTLS the results from the o-suite and control run often differ and the agreement with observations is worse.

Figure 3.2.8.a Daily profiles for ozone from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over East Asia during JJA 2019.
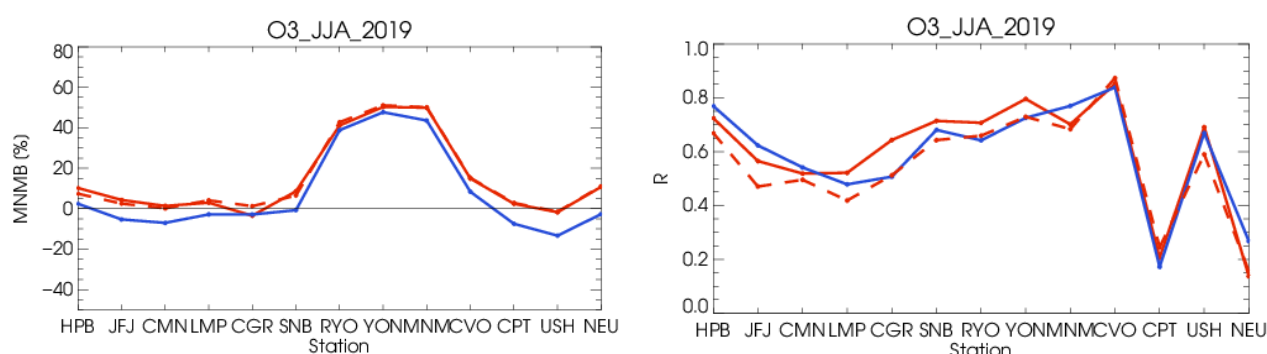
Figure 3.3.1: Modified normalized mean bias in % (left) and correlation coefficient (right) of the NRT model runs compared to observational GAW data in the period June 2019 to August 2019 (o-suite: solid red, D+2: red-dashed, and control: blue).

3.3 Validation with GAW and ESRL-GMD surface observations

For the Near Real Time (NRT) validation, 13 GAW stations and 14 ESRL stations are currently delivering O_3 surface concentrations in NRT, and the data are compared to model results. In the following, a seasonal evaluation of model performance for the 2 NRT runs (o-suite and control) has been carried out for the period from June to August 2019. The latest validation results based on GAW stations and based on ESRL observations can be found on the CAMS website, see section 1, table 1.2.

Modified normalized mean biases in % (left panel) and correlation coefficients (right panel) for different forecasts days (D+2, red-dashed and D+4, red- pointed) with respect to GAW and ESRL observations are shown in Figs. 3.3.1 and 3.3.2. It indicates that MNMBs for both o-suite and control run mostly remain stable up to D+4 (forecast run from 96h to 120h). Correlations between simulated and observed surface ozone values remain almost stable up to D+2 (forecast run from 48h to 72h), but then drop (correlations for D+4 are lower than correlations for D+2 and D+0), see 3.3.2, right graph).

A comparison of the seasonal-mean MNMB over Europe (Fig. 3.3.3) from December 2012 to present shows minimal MNMBs during the winter season and larger biases in other months. Also, on average the MNMB for the o-suite and control shows an improvement over the years. The temporal correlation is consistently better for the control run than for the o-suite, but the o-suite shows strong improvements recently. The GAW results are summarized in Figs 3.3.1 and 3.3.3.

Looking at different regions, for European stations (HPB, JFJ, ZUG, SNB, CMN, LMP, CLM, CGR), observed O_3 surface mixing ratios are very close to the observations. MNMBs are between -3% and 10% for the o-suite and between -7% and 2% for the control run, see Fig. 3.3.1. Correlations for European stations are between 0.51 and 0.72 for the o-suite and between 0.47 and 0.76 for the control run, see Fig. 3.3.1.
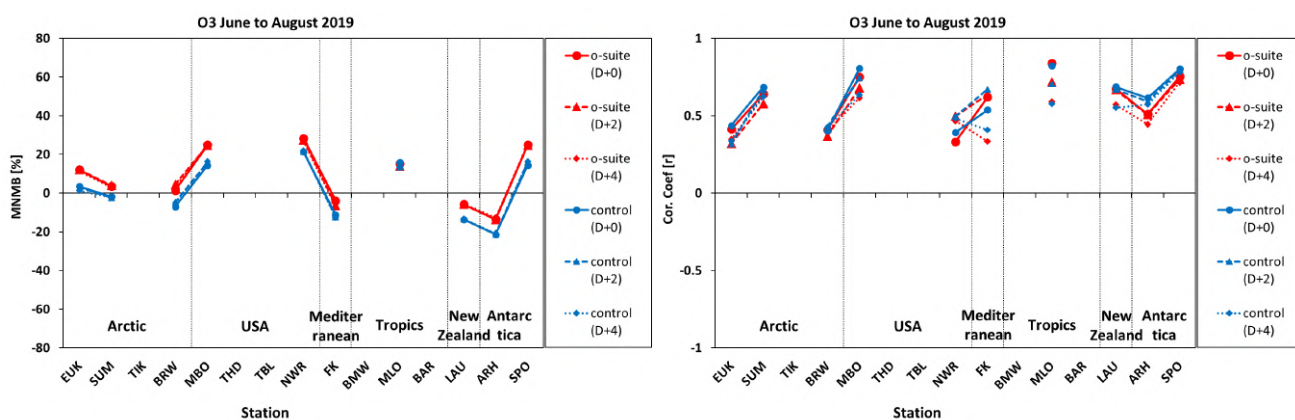
Figure 3.3.2: Modified normalized mean bias in % (left) and correlation coefficient (right) of the NRT forecast runs compared to observational ESRL data in the period June to August 2019. Circles correspond to D+0, triangles to D+2 and rhombs to D+4 metrics respectively.
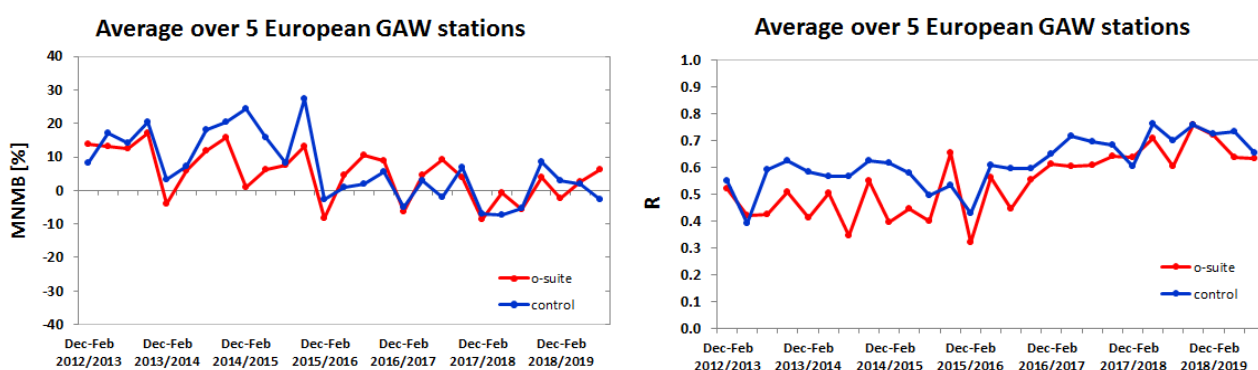

Figure 3.3.3: Long term (Dec. 2012 – August 2019) evolution of seasonal mean MNMB (left) and correlation (right), as averaged over 5 GAW stations in Europe, for o-suite (red) and control (blue).

Over Arctic stations (EUK, BRW and SUM), the o-suite slightly overestimates surface ozone values between 1% (at Point Barrow) and 11% (at Eureka). On the other hand, in the control run O_3 surface mixing ratios are very close to the observations (MNMBs between 3% at EUK and -5% at BRW). Correlations for European stations are between 0.40 and 0.65 for the o-suite and between 0.42 and 0.69 for the control run, see Fig. 3.3.2.

For stations located in Asia (RYO, YON, MNM) both runs strongly overestimate the low observed O_3 mixing ratios in summer with MNMBs between 41% and 50% for the o-suite and between 38% and 47% for the control run, see Fig 3.3.6. Correlation coefficients range between 0.70 and 0.79 for the o-suite and between 0.64 and 0.76 for the control run.

For MBO and NWR USA stations, the observed ozone mixing ratios are overestimated by the o-suite between 24% and 28% respectively. Control run MNMBs are slightly lower ($MNMB_{cnt}$=15% at MBO and 22% at NWR). Correlations between o-suite and observations are between 0.33 (at NWR) and 0.75 (at MBO) for the o-suite and between 0.39 (at NWR) and 0.80 (at MBO) for the control run.
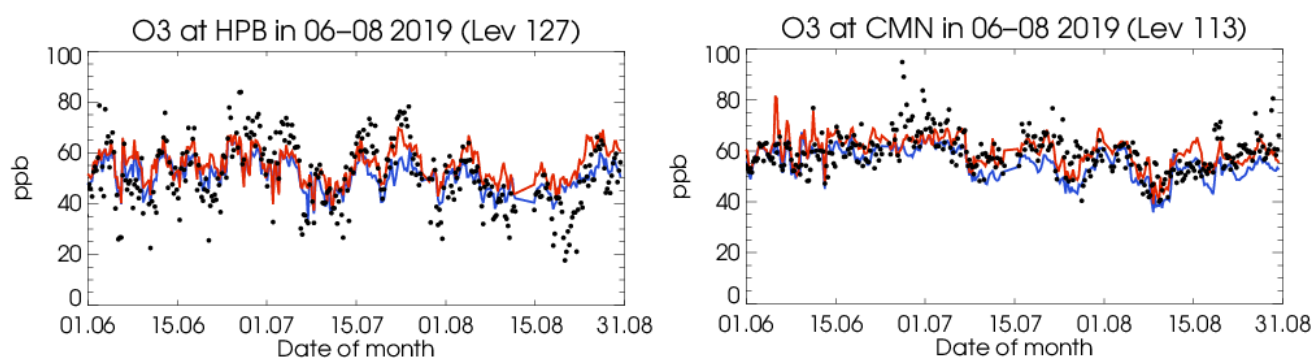
Figure 3.3.4: Time series for the o-suite (red) and control (blue) compared to GAW observations for Hohenpeissenberg (47.8°N, 11.02°E) and Monte Cimone (44.18°N, 10.70°E).
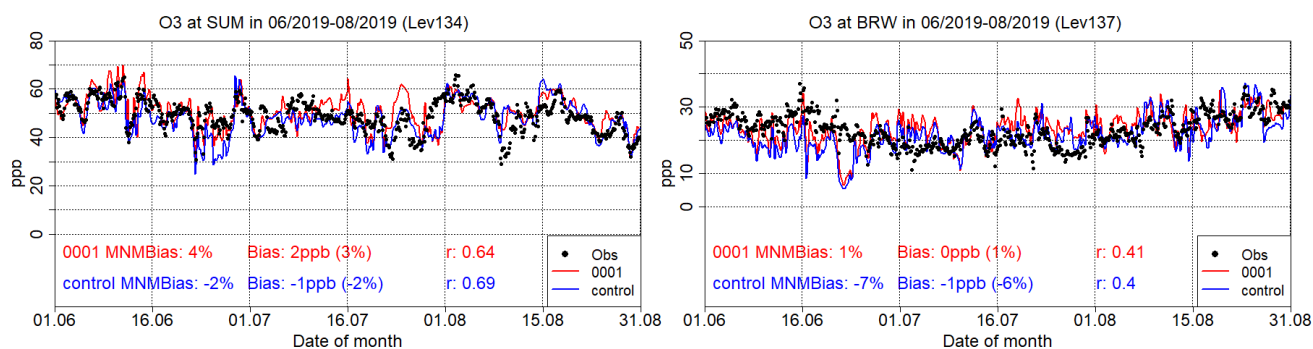
Figure 3.3.5: Time series for the o-suite (red) and control (blue) compared to ESRL observations at Summit, Greenland station (72.57°N, 38.48°W, left) and at Point Barrow, Alaska station (71.32°N, 156.51°W, right)
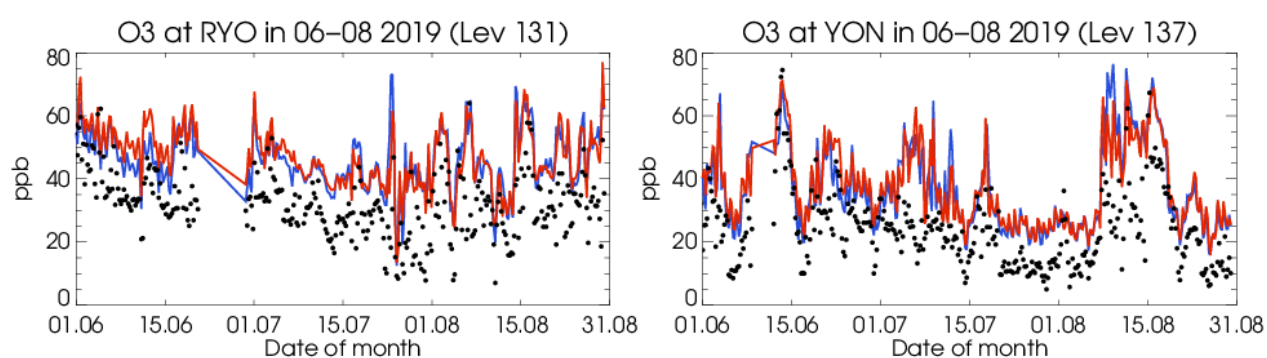

Figure 3.3.6: Time series for the o-suite (red) and control (blue) compared to GAW observations for Ryori (39.03°N, 141.82°E) and Yonagunijima (24.47°N, 123.02°E).

The observed ozone mixing ratios are overestimated by both runs over Mauna Loa (MLO) station in the Tropics (MNMB≈15%). Correlations between simulated and observed surface ozone are high for both the o-suite and the control run (r>0.8).
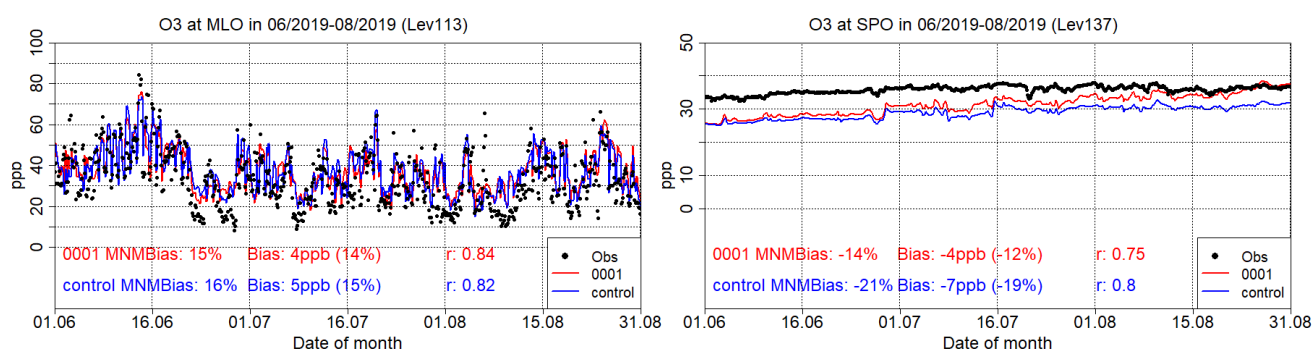
Figure 3.3.7: Time series for the o-suite (red) and control (blue) compared to ESRL observations (black dots) at Mauna Loa, Hawaii station (19.54°N, 155.58°W) and at South Pole, Antarctica station (90.00°S, 24.80°W).
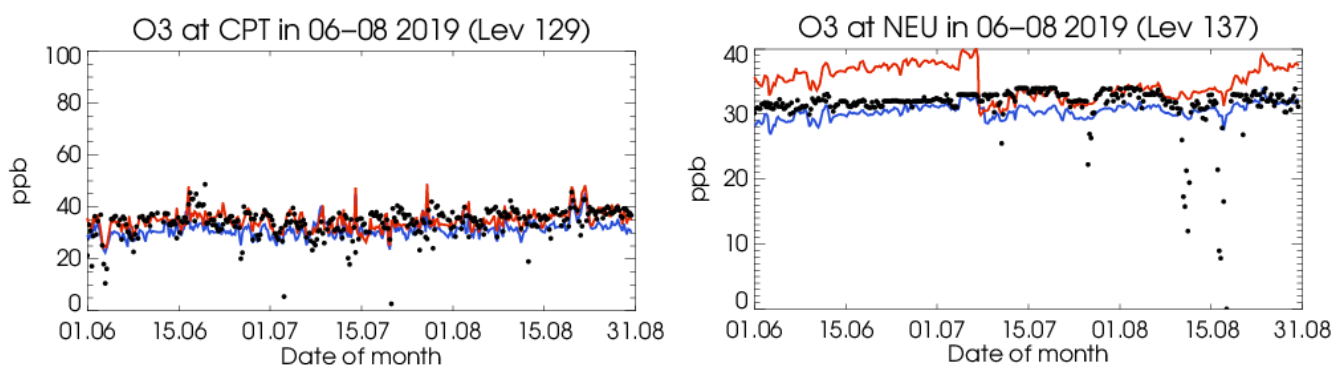
Figure 3.3.8: Time series for the o-suite (red) and control (blue) compared to GAW observations (black dots) at Cape Point (34.55°S, 18.48°W) and GAW observations at Neumayer (70.65°S, 8.25°W).

The O_3 mixing ratios of the southern hemispheric stations (CPT, USH) show MNMBs between -1 and 2% for the o-suite. The control run shows larger underestimations for USH up to -13%, see Fig 3.3.8. Correlation coefficients range around 0.2 and 0.69. At Lauder (LDR) station in New Zealand the o-suite underestimates O_3 mixing ratios by -6%. Again, the control run shows larger underestimation (MNMB=-13%). Correlations between simulated and observed surface ozone values are 0.68 for both runs.

At Arrival Height (ARH) station in Antarctica, the MNMB is -13% for the o-suite and -21% for the control run. Correlation coefficients are 0.45 for the o-suite and 0.50 for the control run. Finally, for South Pole station in Antarctica (SPO), the MNMB is 24% for the o-suite and -14% for the control run. Correlation coefficients 0.75 for the o-suite and 0.80 for the control run. For Neumayer station (NEU) the MNMB is 10% for the o-suite and -2% for the control run.

Especially for the run with data assimilation the change to the new 137 level o-suite is visible, leading to closer correspondence with the measurements from 10th July onwards. Correlation coefficients are low for both o-suites and control.

Table 3.4.1: Coordinates, elevation, corresponding model level (level 137 is the surface level), as well as validation scores (MNMBs and correlations for the period JJA 2019) obtained with the 2 forecast runs (o-suite and control), for each one of the selected Mediterranean stations. MNMBs and correlations with blue denote stations where control run performs better while with red are denoted stations where o-suite performs better.

Station Name	Stat_ID	Lon	Lat	Alt (m)	Level	Distance from the shore (km)	MNMB o-suite	MNMB control	Cor. Coef o-suite	Cor. Coef control
Al Cornocales	ES1648A	-5.66	36.23	189	133	16	38.9	31.2	0.48	0.42
Caravaka	ES1882A	-1.87	38.12	1	137	73	24.6	23.0	-0.01	0.24
Zarra	ES0012R	-1.10	39.08	885	130	70	1.1	-11.3	0.74	0.68
VIillar Del Arzobispo	ES1671A	-0.83	39.71	430	137	48	9.8	4.4	0.44	0.24
Cirat	ES1689A	-0.47	40.05	466	137	37	18.1	15.4	0.42	0.58
Bujaraloz	ES1400A	-0.15	41.51	327	137	60	9.2	1.9	0.60	0.70
Morella	ES1441A	-0.09	40.64	1150	128	51	NA	NA	NA	NA
Bc-La Senia	ES1754A	0.29	40.64	428	137	21	-1.0	-15.2	0.53	0.53
Ay-Gandesa	ES1379A	0.44	41.06	368	136	15	9.0	-4.1	0.68	0.64
Ak-Pardines	ES1310A	2.21	42.31	1226	135	81	18.2	7.2	0.73	0.77
Hospital Joan March	ES1827A	2.69	39.68	172	133	3	15.2	2.0	0.41	0.41
Al-Agullana	ES1201A	2.84	42.39	214	137	25	0.6	-8.2	0.35	0.36
Av-Begur	ES1311A	3.21	41.96	200	132	9	21.3	8.9	0.84	0.73
Plan Aups/Ste Baume	FR03027	5.73	43.34	675	124	21	10.7	0.3	0.60	0.55
Montemonaco	IT1842A	13.34	42.90	1000	127	46	10.2	-0.9	0.35	0.47
Gharb	MT00007	14.20	36.07	114	132	31	-1.6	-9.9	0.62	0.61
Aliartos	GR0001R	23.11	38.37	110	136	18	NA	NA	NA	NA
NEO	-	21.67	37.00	50	137	2	NA	NA	NA	NA
Finokalia	GR0002R	25.67	35.32	250	132	4	-3.9	-11.2	0.62	0.54
Agia Marina	CY0002R	33.06	35.04	532	133	14	6.1	-1.4	0.74	0.76

3.4 Validation with AirBase observations in Mediterranean

The surface ozone validation analysis over the Mediterranean is based on an evaluation against station observations from the Airbase Network (http://acm.eionet.europa.eu/databases/airbase/). In addition, 1 station from the Department of Labour Inspection - Ministry of Labour and Social Insurance, of Cyprus (http://www.airquality.dli.mlsi.gov.cy/) is used in the validation analysis. For the validation analysis, stations in the Mediterranean located within about 100 km from the shoreline of the Mediterranean shore are used. Table 3.4.1 shows the names, coordinates, elevation and the MNMBs and correlations obtained with the 2 forecast runs (o-suite and control). It indicates that the variance explained by each station of both the o-suite and control is high and correlations are highly significant over Western, Central and Eastern Mediterranean. It should be noted that the control run reproduces slightly better than the o-suite run the surface ozone day to day variability over almost all the Mediterranean stations (see Table 3.4.1, exception is the Agia marina Cyprus station).

In terms of biases, the o-suite mostly overestimates surface ozone values and its MNMBs vary between -1% and +40% depending on the stations over the Mediterranean shore of Spain (average MNMB for the 13 Spain Mediterranean station is 14%). For the control run MNMBs are on average 9.5% and thus lower than the o-suite's MNMBs. Over the stations Plan Aups/Ste Baume in France and Montemonaco in Italy the o-suite overestimates surface ozone concentrations by 10%. Over Gharb station in Malta the o-suite underestimates surface ozone values by -2% and the control run by -10%. Over Finokalia station in Crete the o-suite underestimates surface ozone by -4% while the control run underestimates it by -11%. Finally, over Agia Marina in Cyprus the o-suite overestimates surface ozone values by 6% while the control run has almost zero MNMB.
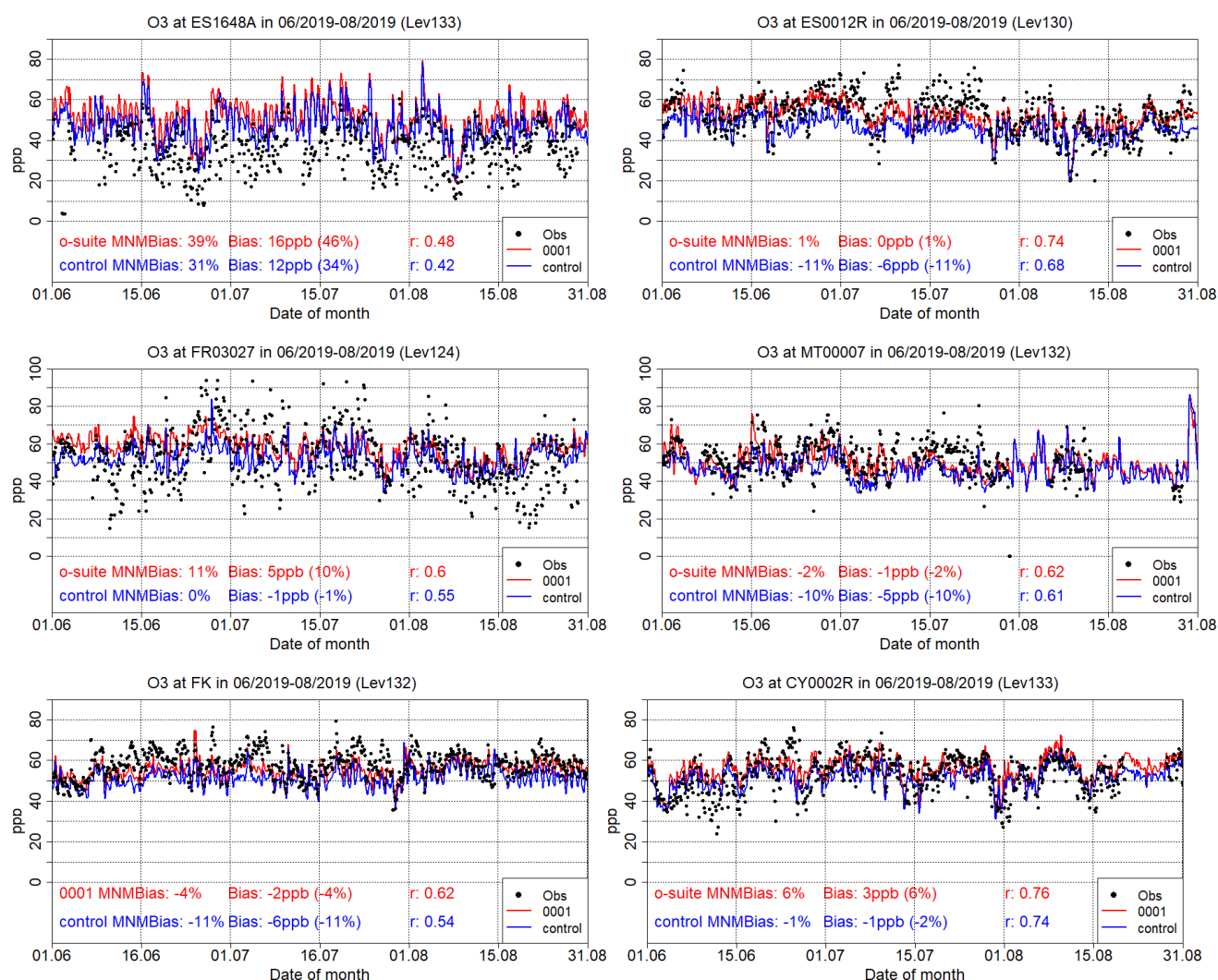
Figure 3.4.1: Time series for the o-suite (red) and Control (blue) compared to Airbase observations at Al Cornocales, Spain station (36.23°N, 5.66 °W, top left), at Zarra, Spain station (39.08°N, 1.10°W, top right), at Plan Aups/Ste Baume, France station (43.34°N, 5.73°E, center left), at Gharb, Malta station (36.07°N, 14.20°E, center right at Finokalia, Crete Greece station (35.32°N, 25.67°E, bottom left) and compared to observations provided by the Department of Labour Inspection - Ministry of Labour and Social Insurance of Cyprus) at Agia Marina, Cyprus station (35.04°N, 33.06 °E, low right).

The spatial distribution of MNMBs and the correlation coefficients of the o-suite over the Mediterranean are shown in 3.4.2, where it is evident that correlations over the entire Mediterranean from Gibraltar to Cyprus are highly significant (exception is Caravaca, Spain station). It is also evident that the CAMS NRT runs have a better performance over Central and eastern Mediterranean compared to the Mediterranean shore of Spain in terms of biases.
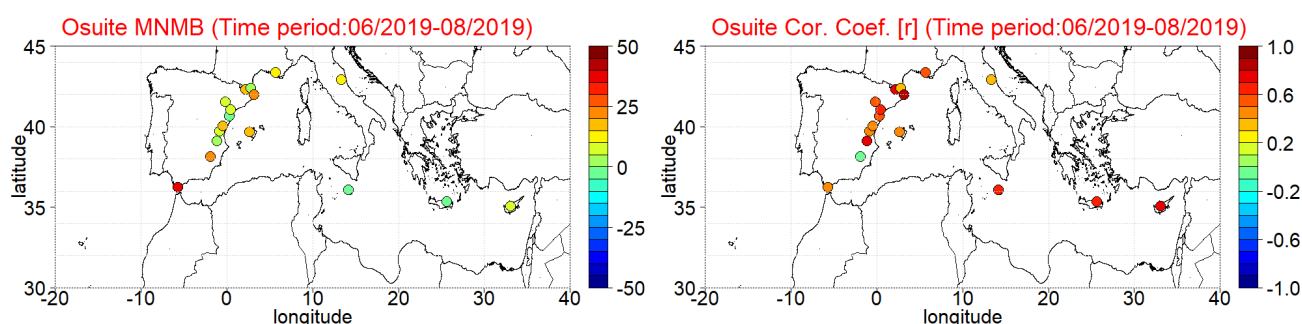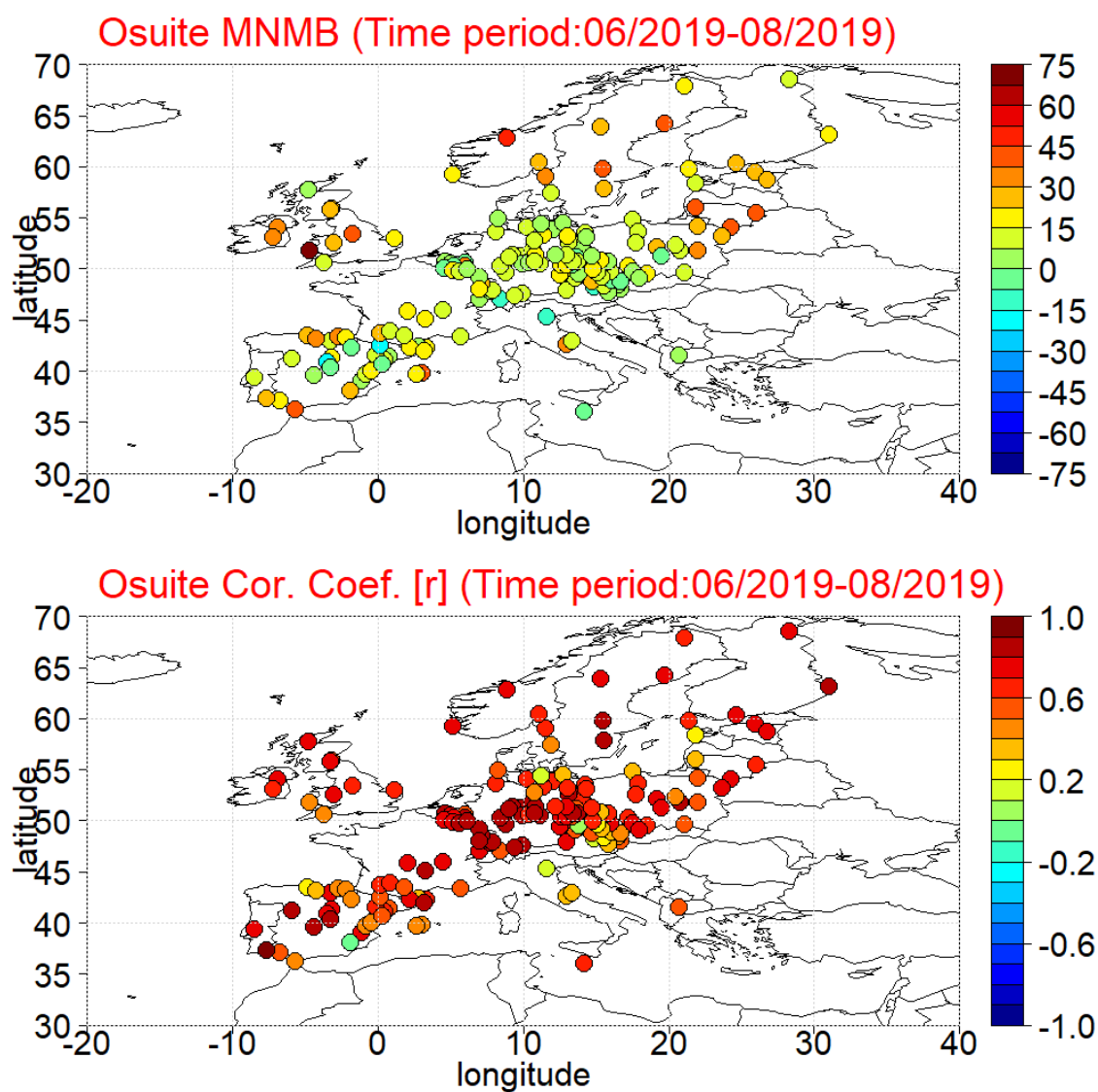
Figure 3.4.2: Spatial distribution of MNMB in % (left) and correlation coefficient (right) of the o-suite run compared to observational data during the period from 1 June 2019 to 31 August 2019.

3.5 Validation with AirBase observations over Europe

The surface ozone validation analysis over Europe is based on an evaluation against Background rural Classes 1-2 O_3 July-Peuch classification station observations from Airbase Network (http://acm.eionet.europa.eu/databases/airbase/). The spatial distribution of MNMBs and the correlation coefficients of the o-suite over Europe are shown in 3.5.1, where it is evident that correlations over the majority of European AirBase stations (with a very few exceptions) are highly significant. It is also evident that the CAMS NRT runs reproduce well the surface ozone mean concentrations over central Europe (depending on the station MNMBs vary from -10% to +20%). It is also evident that over the Iberian Peninsula, Great Britain, Poland, Baltic countries and Scandinavia the o-suite mostly overestimates surface ozone values with maximum MNMBs values (exceeding 50% in few stations) observed in UK, Sweden and Lithuania. The above-mentioned findings concerning CAMS NRT runs biases and correlations are also observed in individual time series at selected stations plotted in Figure 3.5.2. From this time series and the plotted validation metrics is also evident that control run surface ozone concentrations are 4-5 ppb lower than the o-suite values resulting in most cases a closer to zero bias. From this time series it is also evident that during the end of June and the end of July high ozone episodes (with observed surface ozone values up to 100 ppb) occurred over Western Europe mainly over France, Germany and Belgium. It can also clearly be seen that although the o-suite mostly overestimates surface ozone values during summer it underestimates them during the high ozone episodes. More details and validation metrics concerning this summer 2019 high ozone episodes are given in a specific section in the event studies part of this report.

Figure 3.5.1: Spatial distribution of MNMB in % (left) and correlation coefficient (right) of the o-suite run compared to observational data during the period from 1 June 2019 to 31 August 2019.
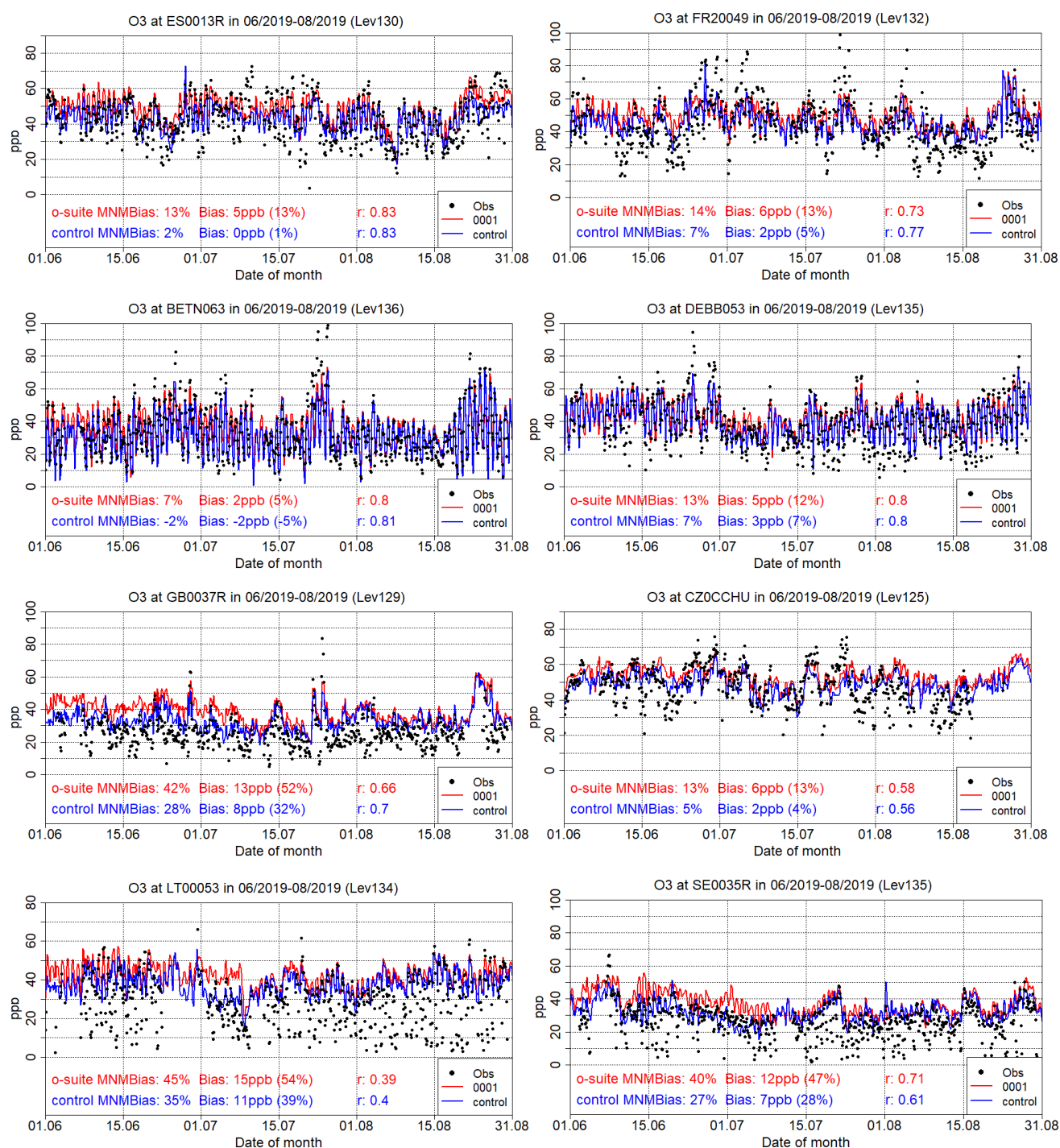
Figure 3.5.2: Time series for the o-suite (red) and Control (blue) compared to Airbase observations at Al Penausende, Spain station (41.24°N, 5.90 °W, 1st row left), at Haut Beaujolais, France station (45.96°N, 4.47°E, 1st row right), at Corroy L.G., Belgium Station (50.67°N, 4.67°E, 2nd row left), at Hasenholz, Germany (52.56°N, 14.02°E, 2nd row rigth), at Ladybower, Great Britain station (53.40°N, 1.75°W, 3nd row left), at LdGajewWIOSAGajew, Poland station (52.14°N, 19.23°E 3nd row rigth), at Zemaitija, Lithuania station (56.01°N, 21.89°E, 4nd row left) and at Vindeln, Sweden station (64.25°N, 19.77°E, 4nd row right).

3.6 Validation with IASOA surface observations

The CAMS results were compared to surface O_3 observations from the Villum Research Station, Station Nord in north Greenland (81.6°N 16.7°W), and Zeppelin Mountain, Svalbard (78.9°N 11.9°E) from the IASOA network (Fig. 3.5.1).

The data from Svalbard and VRS are covering the period from December 2014 to August 2019. The model simulations do not capture ozone depletion events in March – June in 2015 – 2019 during spring at any of the sites. These events are related to halogen chemistry reactions that are not represented in the model simulations. The simulations are on average in good agreement with the observations apart from the spring depletion events.

For the period June – August 2019 the measurements are not quality controlled. The model simulations underestimate the measured O_3 concentrations at the two sites resulting in a negative bias of 2% - 14% for the o-suite and 2% - 13% for the control run (Table 3.6.1) like previous years. The control run performs better than the o-suite in terms of the correlation; r = 0.38 – 0.58 for the o-suite compared to r = 0.41 – 0.67 for the control run.

Table 3.6.1. Modified Normalised Mean Bias (MNMB) and correlation coefficient (r) of the O-suite and the Control simulations for the sites Svalbard, and Villum Research Station (VRS) for the period June – August 2019.

		MNMB	R
Svalbard	o-suite	-0.02	0.58
	control	-0.02	0.67
VRS	o-suite	-0.14	0.38
	control	-0.13	0.41
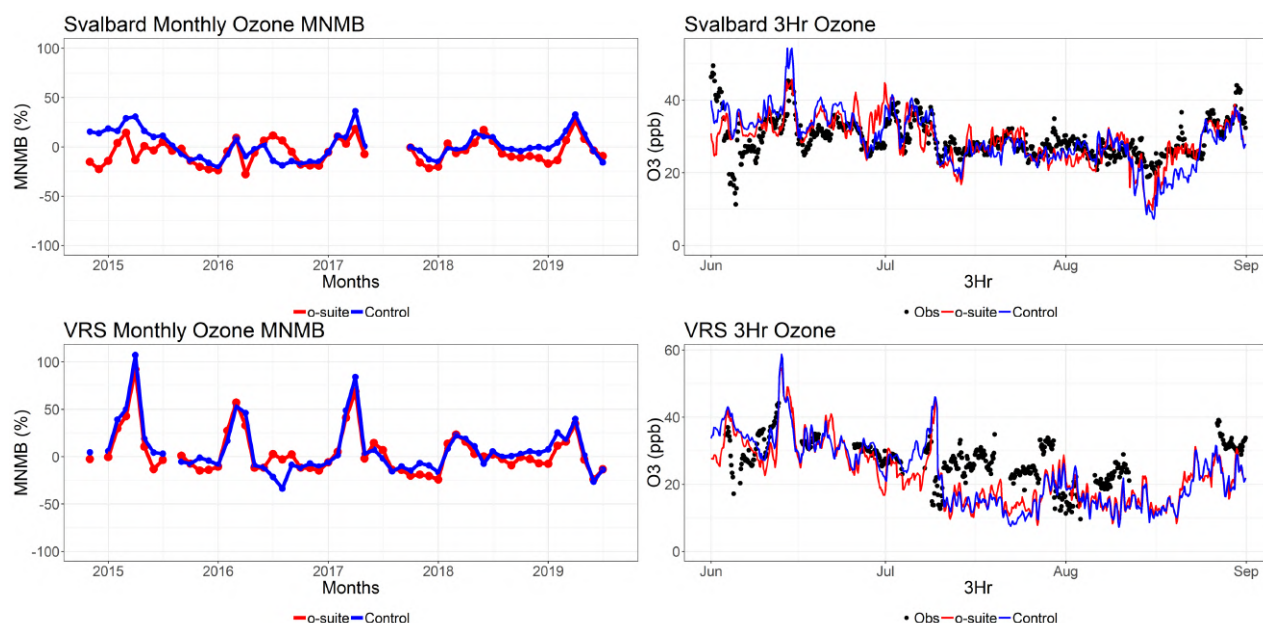
Figure 3.6.1: Time series for o-suite (red) and Control (blue) compared to observations (black dots) at, Svalbard (top row), and the Villum Research Station, Station Nord, Greenland (bottom row) MNMB for the full period (left) and concentrations for March-May (right).

3.7 Validation with IASI data

Ozone total columns from the o-suite and control run are compared with IASI Metop-A version v20151001 daytime only satellite observations (Clerbaux et al., 2009). For the comparison with the IASI data, the vertically integrated model O_3 data were transformed using IASI averaging kernels (Rodgers, 2000).

The global distribution of the O_3 total column obtained from IASI, as well as the relative difference between the model runs and IASI, are shown in Fig. 3.7.1 for August 2019. Satellite data shows high O_3 over the southern mid-latitudes over the Indian Ocean and Norwegian Sea area and low values over the ocean south of Patagonia. The o-suite run captures both, high and low O_3 values relatively well and is in good agreement with the observations, showing MNMBs within 5%. The control run is mainly positively biased (up to 20% over the high northern latitudes). Underestimation can be seen over the biomass burning areas in Africa (within 10%). The forecast day 4 is almost similar to the forecast day 0. Note, that the IASI sensitivity is the lowest over the cold surfaces of Antarctica and Greenland (especially during March-April-May season) where IASI O_3 values are positively biased by up to 20%.
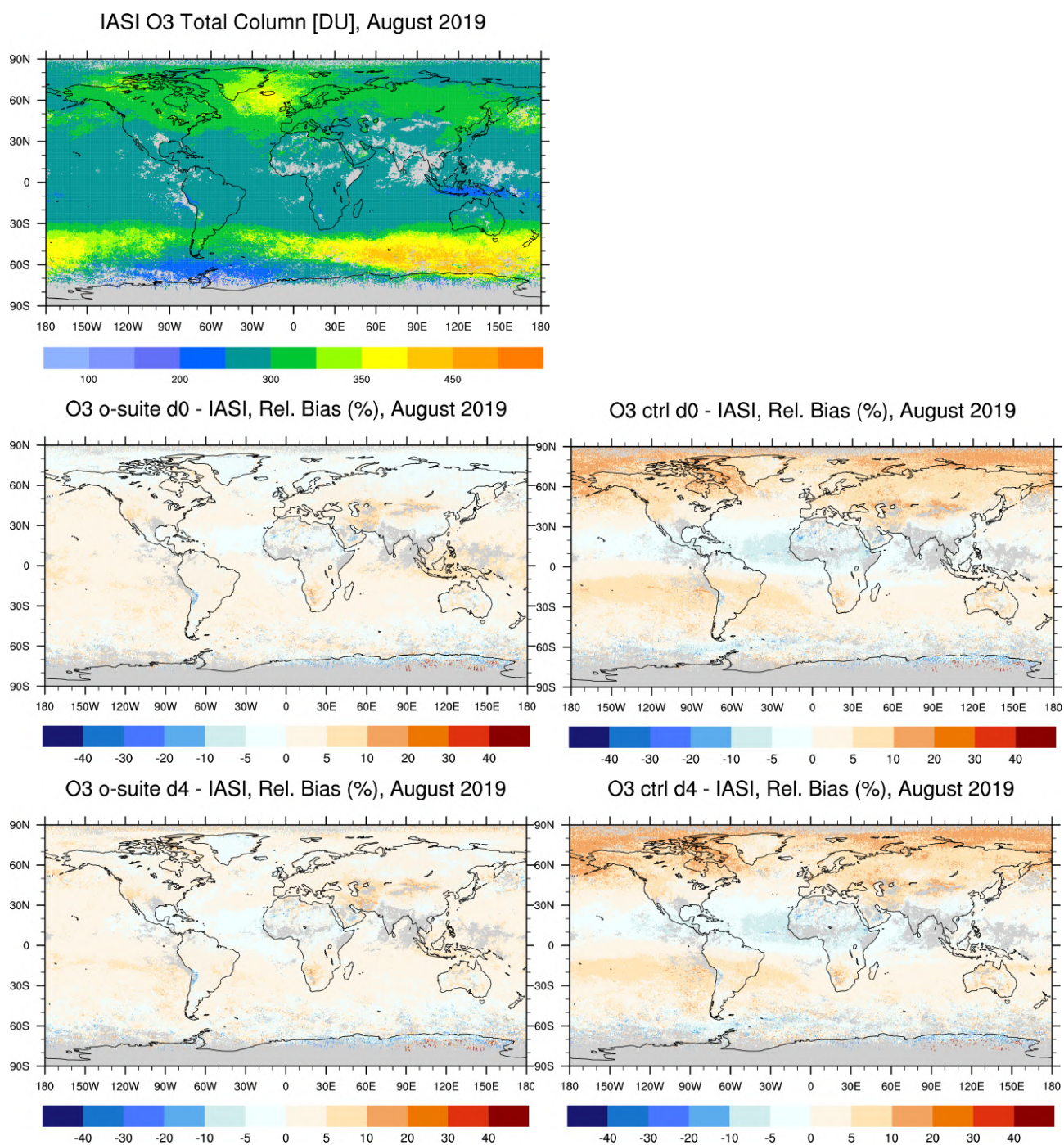
Figure 3.7.1: O$_3$ total column for IASI satellite observations (top) and relative difference between the model runs and IASI for August 2019: o-suite day 0 and day 4 (left), control run day 0 and day 4 (right). Grey colour indicates missing values.

4. Carbon monoxide

4.1 Validation with Global Atmosphere Watch (GAW) Surface Observations

For the Near-Real-Time (NRT) validation, 8 GAW stations have delivered CO surface mixing ratios in NRT and data is compared to model results as described in Eskes et al. (2018) and is used for CAMS model evaluation for June 2019– August 2019. The latest validation results can be found on the CAMS website, see section 1.

For stations in the Northern Hemisphere, both runs mostly show slightly negative MNMBs. For CPT station in the Southern Hemisphere, especially the control run shows a slight positive offset.

For most stations, the MNMBs and correlation coefficients indicate that the forecast remains stable for the D+2.
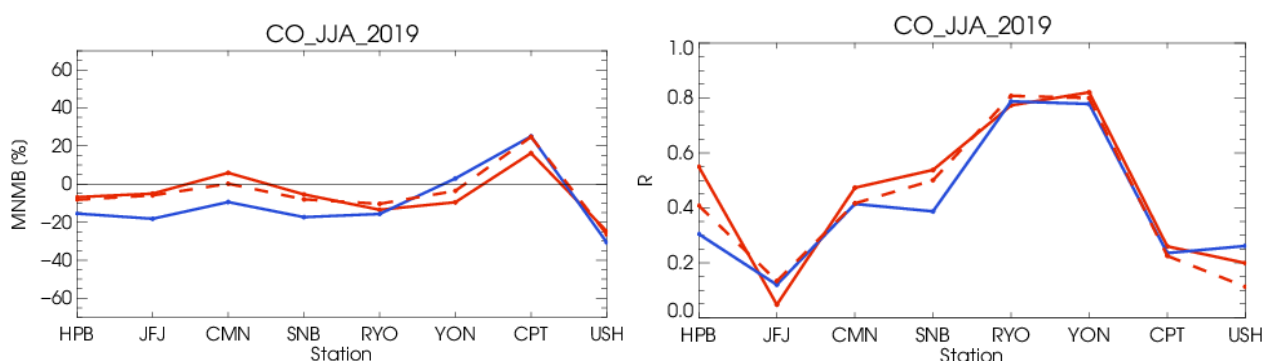
Figure 4.1.1: Modified normalized mean bias in % (left) and correlation coefficient (bottom right) of the NRT model runs compared to observational GAW data in the period March 2019 to August 2019 (o-suite: solid red, D+2: red-dashed, and control: blue). For a list of stations, see Eskes et al. (2018a).
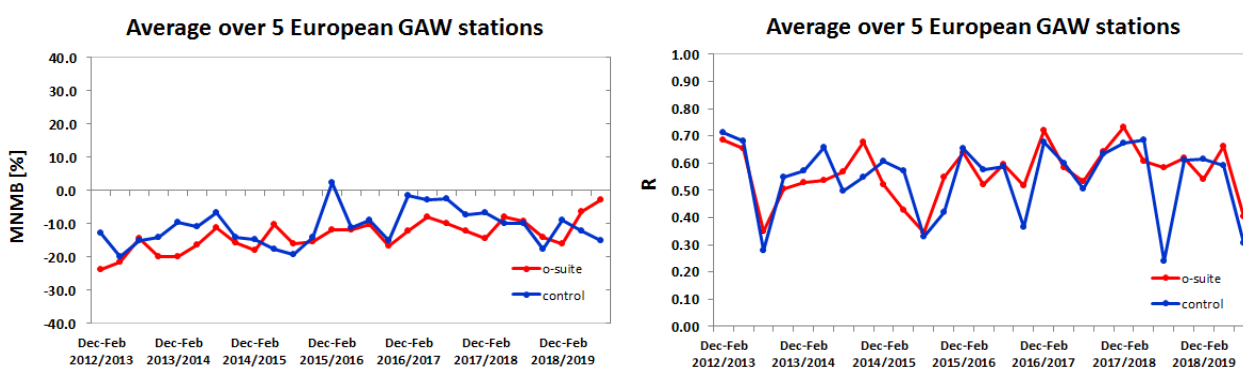
Figure 4.1.2: Long term (Dec. 2012 – August 2019) evolution of seasonal mean MNMB (left) and correlation (right), as averaged over 5 GAW stations in Europe, for o-suite (red) and control (blue).
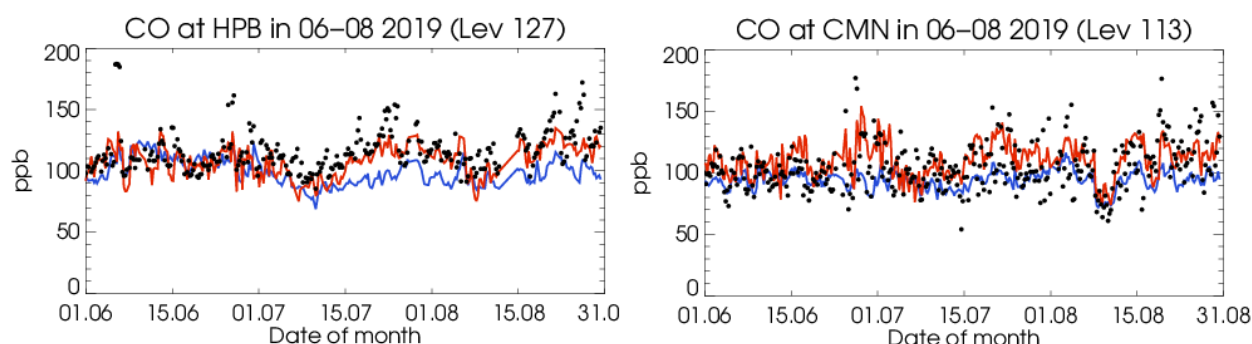
Figure 4.1.3: Time series for the o-suite (red) and control (blue) compared to GAW observations at Hohenpeissenberg (47.8°N, 11.02°E) and Monte Cimone (44.18°N, 10.7°E).
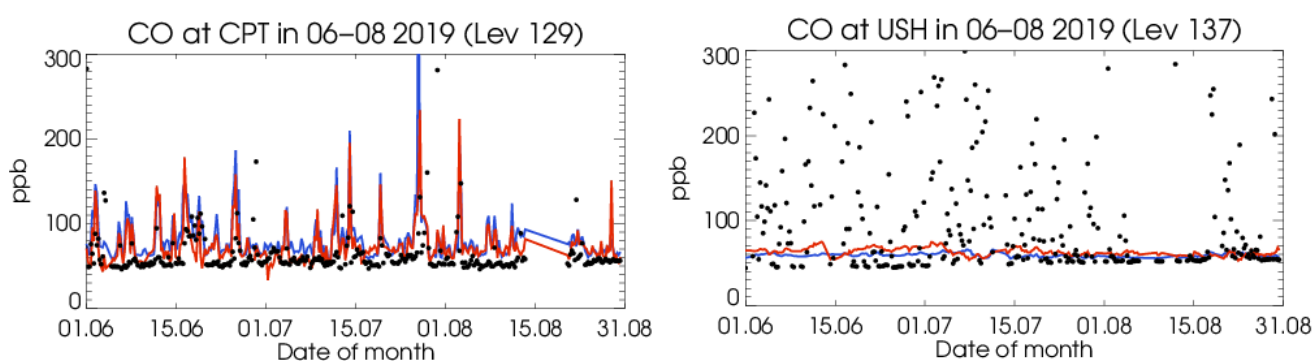
Figure 4.1.4: Time series of the o-suite (red) and control (blue) compared to GAW observations for Cape Point (34.35°S, 18.5°E) and Ushuaia (-54.85°N, -68.32°W).

A comparison of the seasonal-mean MNMB over Europe (Fig. 4.1.2) from December 2012 to present shows a slowly improving MNMB from about -20% in 2013 to about -10% for more recent periods. Temporal correlation remains relatively constant at r=0.6 on average, except for the quarter JJA in 2018 and 2019, where the correlation of the control run drops to 0.24, and 0.31, respectively.

For European stations, the o-suite shows an underestimation of observed CO mixing ratios, with MNMBs between 5% and -6%. The control shows more negative MNMBs between -9% and -18%. Correlation coefficients are low between 0 (0.1) and 0.53 (0.41) for the o-suite (control run).

The comparison for the new 137 level run seems to lead to a larger negative offset in the control run from July 10th onwards (Fig. 4.1.3.). Note that the model level selected for the comparisons changed in moving from 60 to 137 vertical levels.

For the two stations in the Southern Mid-latitudes (CPT and USH) the control partly shows an overestimation of CO (CPT) with MNMBs between -31% (USH) and 25%, (Fig.4.1.4.) MNMBs of the o-suite are smaller amounting between -24% (USH) and 16%%.

For stations in Asia (RYO, YON) both runs mostly show a good correspondence with the observations with MNMBs between -13% and -9% for the o-suite and between 2% and -15% for the control run, see Fig. 4.1.5. Correlation coefficients are high between 0.77 and 0.82 for the o-suite and for the control run.
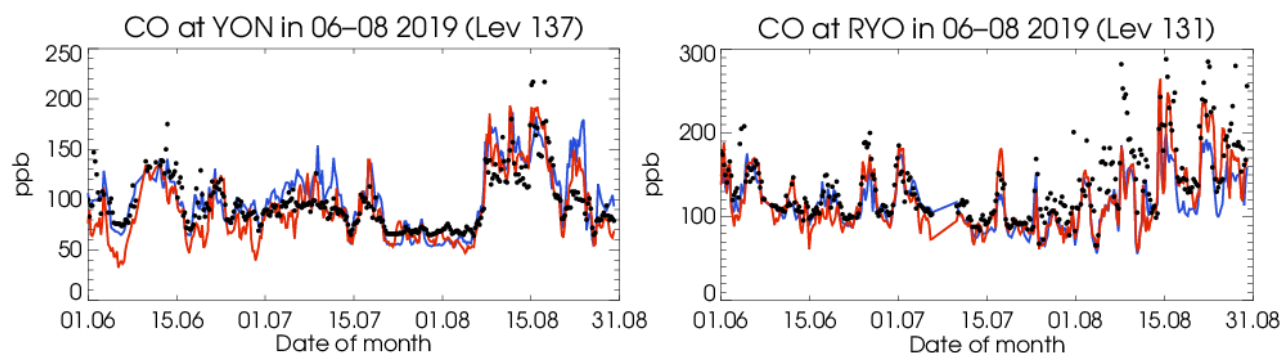
Figure 4.1.5: Time series for the o-suite (red) and control (blue) compared to GAW observations for Yonagunijima (24.47°N, 132.02°E) and Ryori (39.03°N, 141.82°E).

4.2 Validation with IAGOS Data

Like for ozone, continuous time series of CO are only available at Frankfurt during JJA 2019 (Fig. 4.2.1 and 4.2.2). CO is mostly underestimated by both o-suite and associated control run and the largest bias is generally found in the lowest layers. While the performances of the two runs are more similar in the lowest layers, the o-suite performs clearly better than control run in the free troposphere. When the period before and after the upgrade are compared, it seems that in the results obtained with the new assimilated run the bias is improved in the low to mid-troposphere especially in the lowest layers. For the control run the results remain very similar before and after the upgrade.

Several peaks of CO are observed in the surface and boundary layer in the last week of June and July (Fig. 5.2.1), corresponding to the heat wave periods which are correlated with high ozone episodes (see IAGOS ozone section). The individual profiles on several dates during these periods are presented in Fig. 4.2.3a-b. For 2 profiles measured on 26 June, CO reaches 300 ppbv near the surface. In the following days CO values increase in the boundary layer. On 30 June the profile is nearly constant from the surface to the boundary layer top with concentrations of about 200 ppbv. At 16:58 UTC, CO values are of about 150 up to 4000 m. For these profiles CO is largely underestimated by both runs with a slightly better performance from the o-suite. In the lower part of the free troposphere the o-suite performs better, whereas the bias of the control run remains large. Regarding the heat wave of late July, the similar results are obtained with large underestimations from both models as shown on the profiles of 25/07 and 27/07 (Fig. 4.2.3.b).

On the 18 July a strong maximum of CO is observed in two profiles near 4500 m with a magnitude slightly beyond 200 ppbv. This maximum is related to the long-range transport of pollution emitted from forest fires in North America. This maximum is very well reproduced by both models. On 7 and 10 August other maxima of about 200 ppbv are observed near 4000 m (Fig. 4.2.3b). This maximum is the result of the transport of emissions from fires in Siberia to Europe that have started during July. Only one of the plumes is reproduced by the o-suite and the magnitude is largely underestimated while control run is not showing any enhancement (Fig. 4.2.3b).
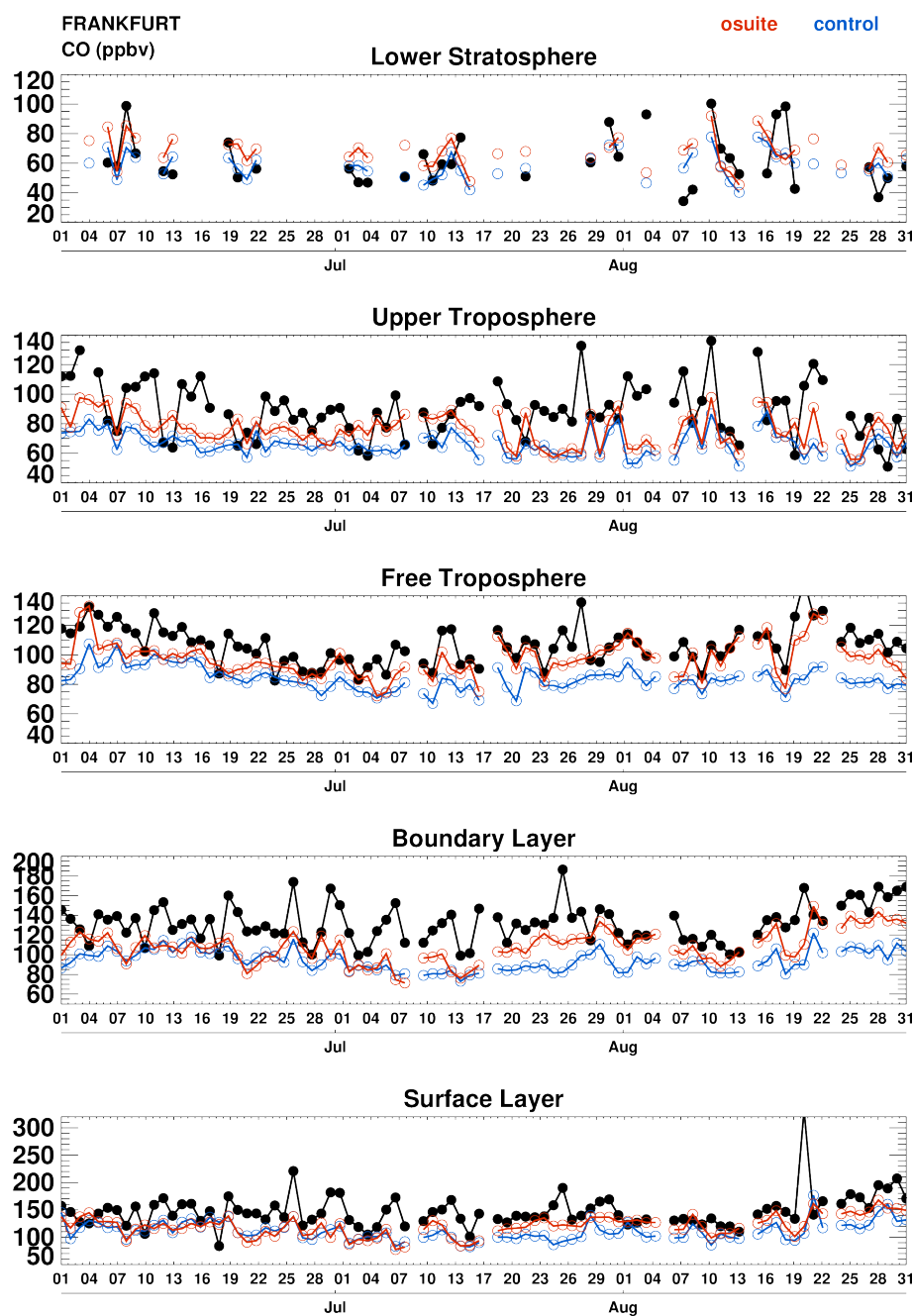
Figure 4.2.1 Time series of daily mean CO over Frankfurt during JJA 2019 for 5 layers: Surface Layer, Boundary Layer, Free Troposphere, Upper Troposphere and Lower Stratosphere. The o-suite is shown in red and associated control run in blue.
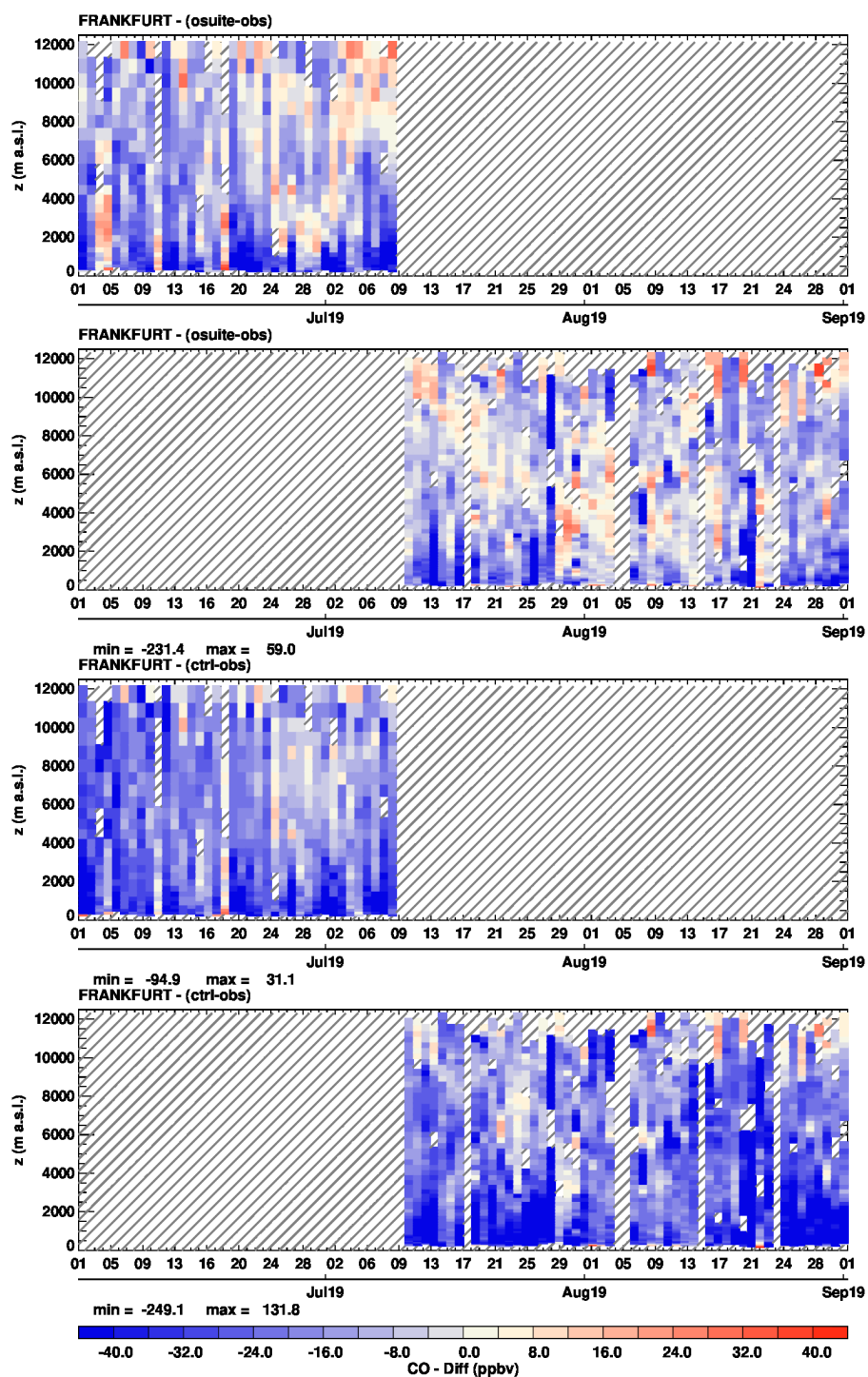
Figure 4.2.2. Time series of the absolute differences (model - observations) in daily profiles for CO over Frankfurt during JJA 2019. Two upper panels correspond to o-suite (before and after upgrade) and two lower panels to control run (before and after upgrade).
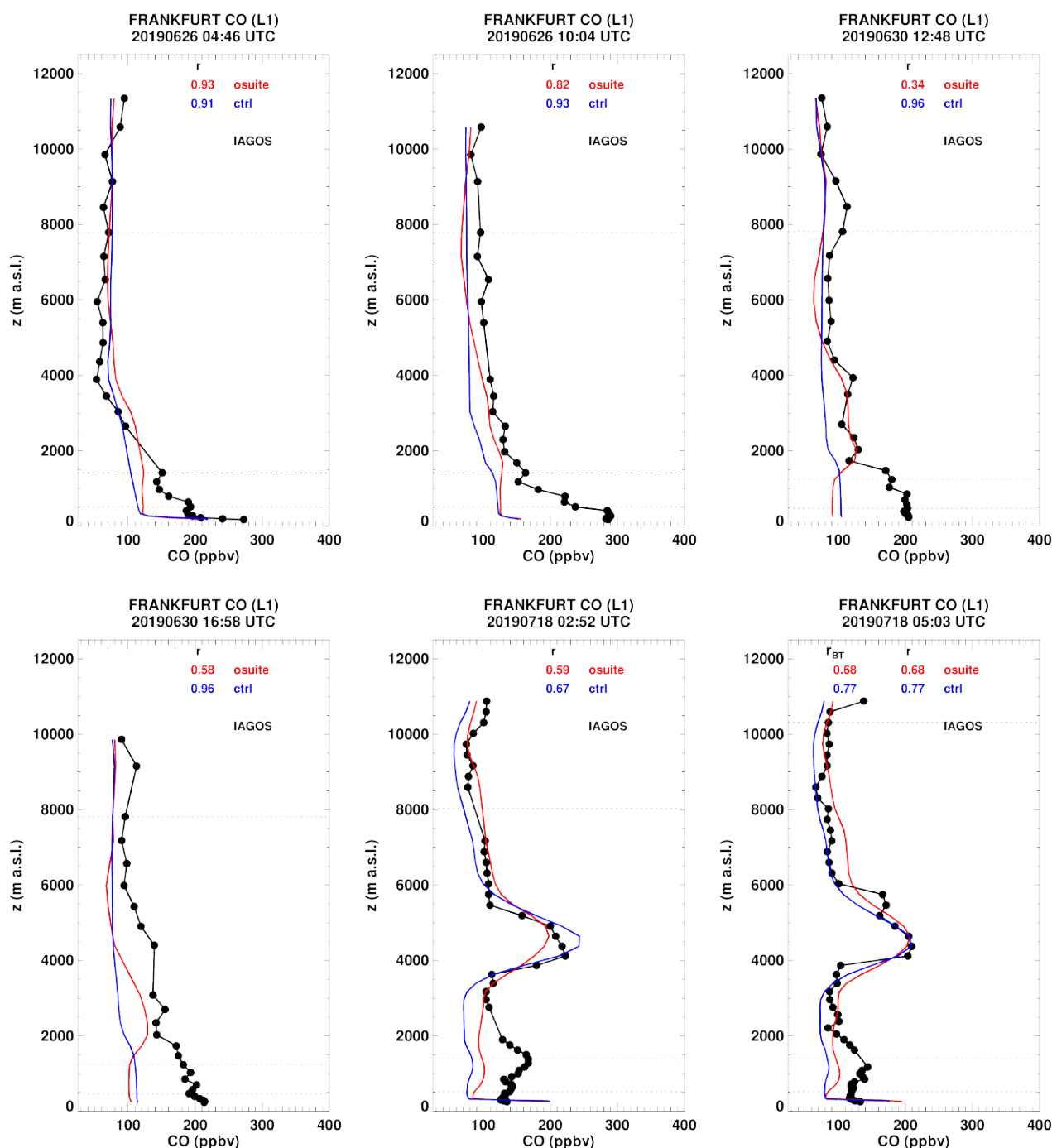
Figure 4.2.3a. Daily profile for CO from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over the Europe during JJA 2019.

FRANKFURT CO (L1)
20190725 06:12 UTC

FRANKFURT CO (L1)
20190727 04:26 UTC

FRANKFURT CO (L1)
20190730 09:23 UTC

FRANKFURT CO (L1)
20190807 08:34 UTC

FRANKFURT CO (L1)
20190807 12:38 UTC

FRANKFURT CO (L1)
20190810 08:41 UTC
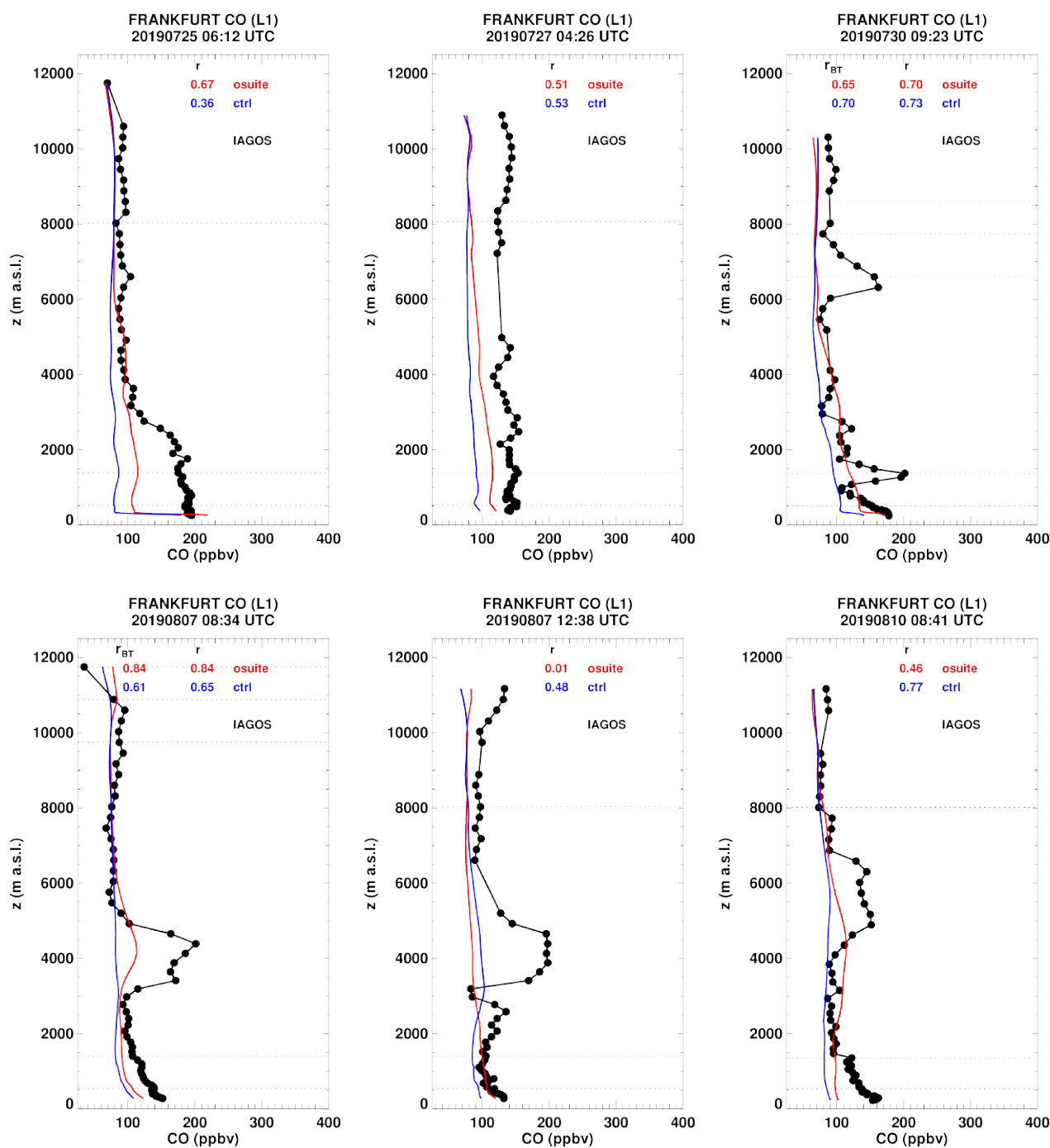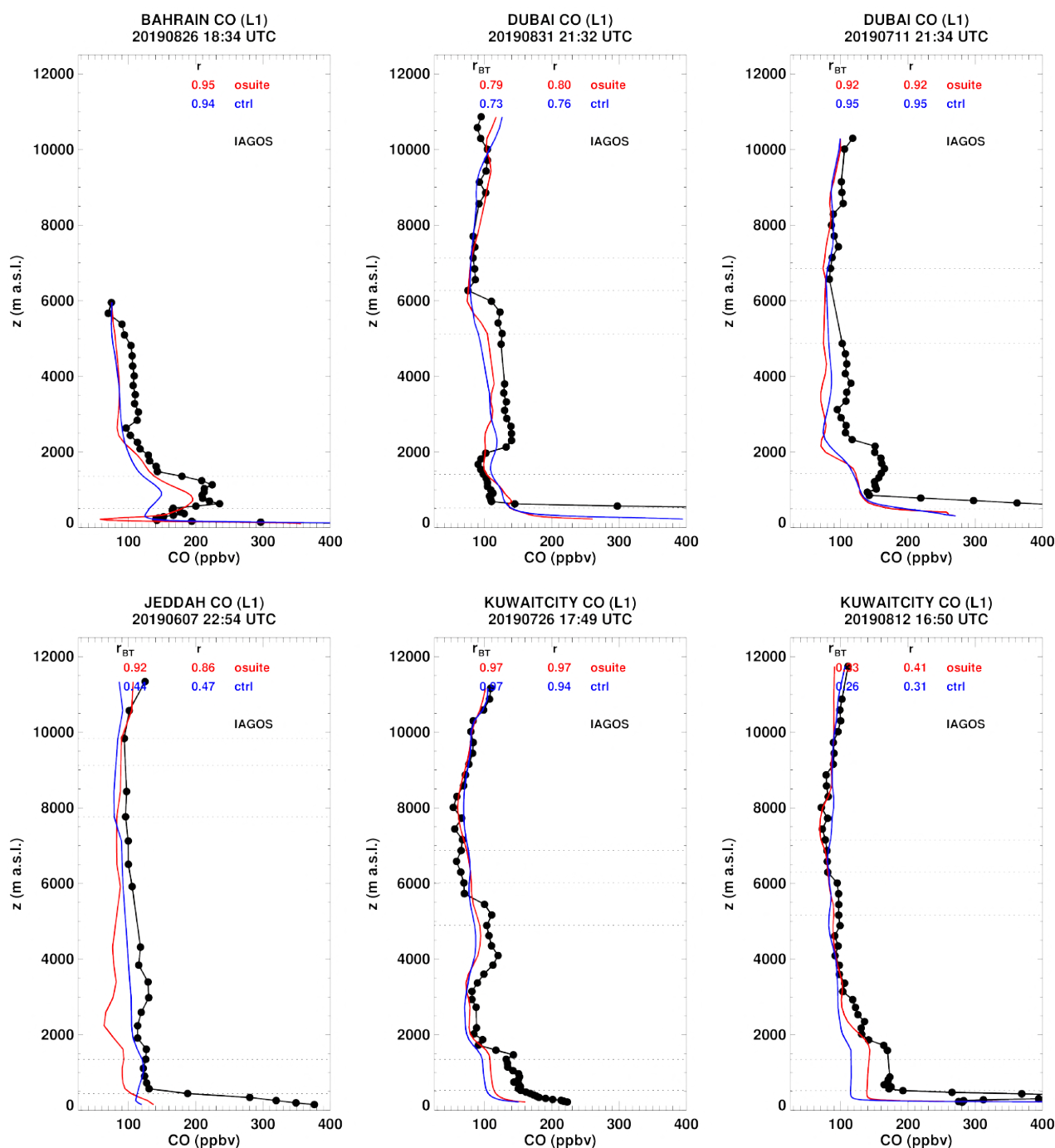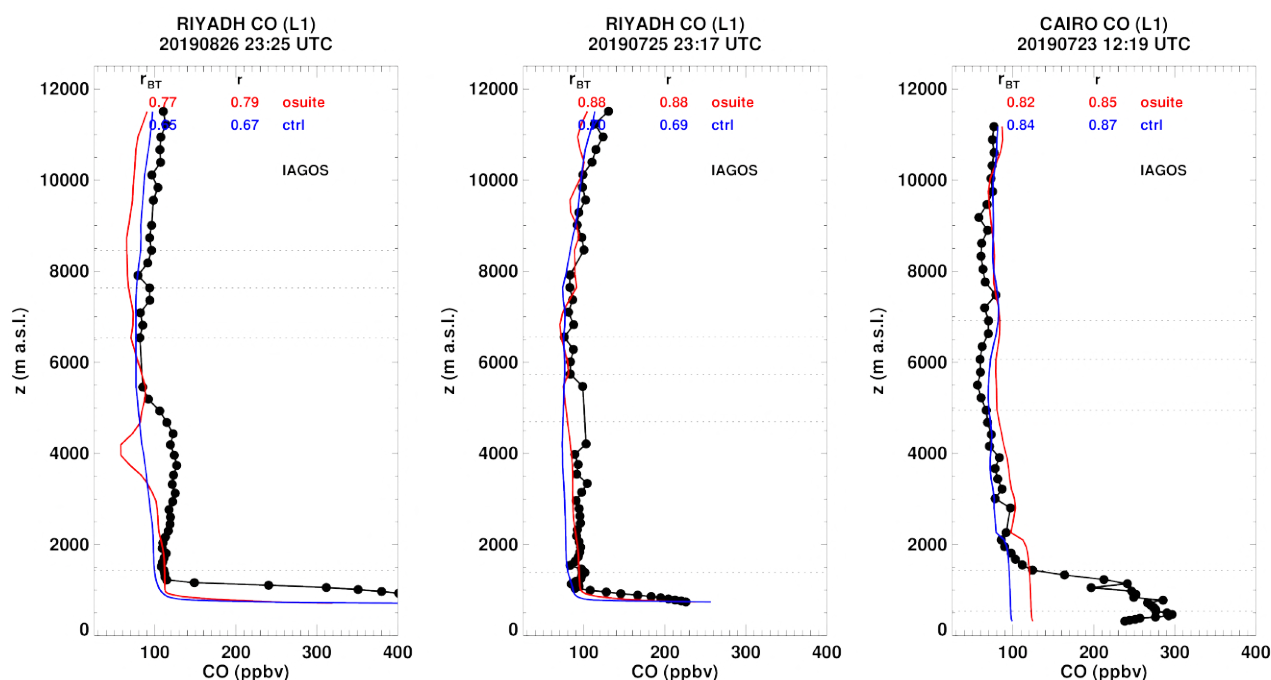
Figure 4.2.3b. Daily profile for CO from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over the Europe during JJA 2019.
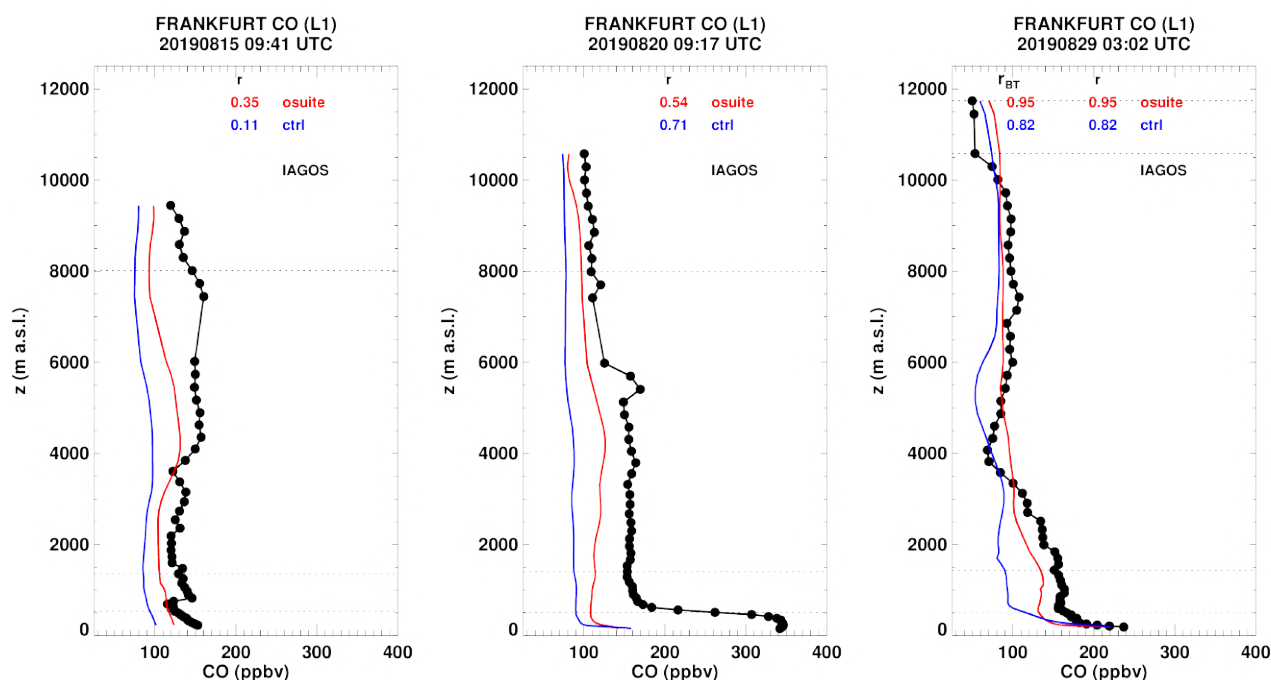
Figure 4.2.3c. Daily profile for CO from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over the Europe during JJA 2019.

In the second half of August high CO values are also observed in the low troposphere and in the free troposphere as shown in both timeseries (Fig. 4.2.1) and individual profiles (Fig. 4.2.3c). A major peak is found in the surface layer on 20 August with CO mixing ratio reaching 350 ppbv on the profile at 9:17 UTC. CO values are also high in the free troposphere with a constant value of about 170 ppbv. The two models behave similarly in the lowest layers and show large underestimations. In the free troposphere the CO increase is detected only by the o-suite although the values are underestimated. On the profile of 15 August, CO values in the free troposphere present similar values as those observed on 20 August and the results from the models are similar to those obtained for day 15. These high CO episodes in the free troposphere are likely related to the contribution of forest fire emissions from Siberia that are transported to Europe. On 29 August, CO reaches 250 ppbv in the surface layer and about 150 ppbv in the boundary layer and the o-suite present a good agreement with observations with slight underestimations while the bias is much larger for the control run.

Middle East

Several profiles are available over the Middle East at the airports of Bahrain, Dubai, Jeddah, Kuwait City, Riyadh and Cairo (Fig. 4.2.4a-b). High mixing ratios of CO reaching 400 ppbv in the surface layer are observed for some profiles. These values are often well reproduced by the models which behave similarly. In the boundary layer the values are often underestimated. Some profiles show small maxima of CO in the free troposphere (about 130 ppbv) like at Kuwait City on 31 August, at Dubai on 26 July, and Riyadh on 26 August. These features are detected by the models with a slightly better performance from the o-suite (smaller underestimations) than from control run. In the UTLS both runs agree well with observations.
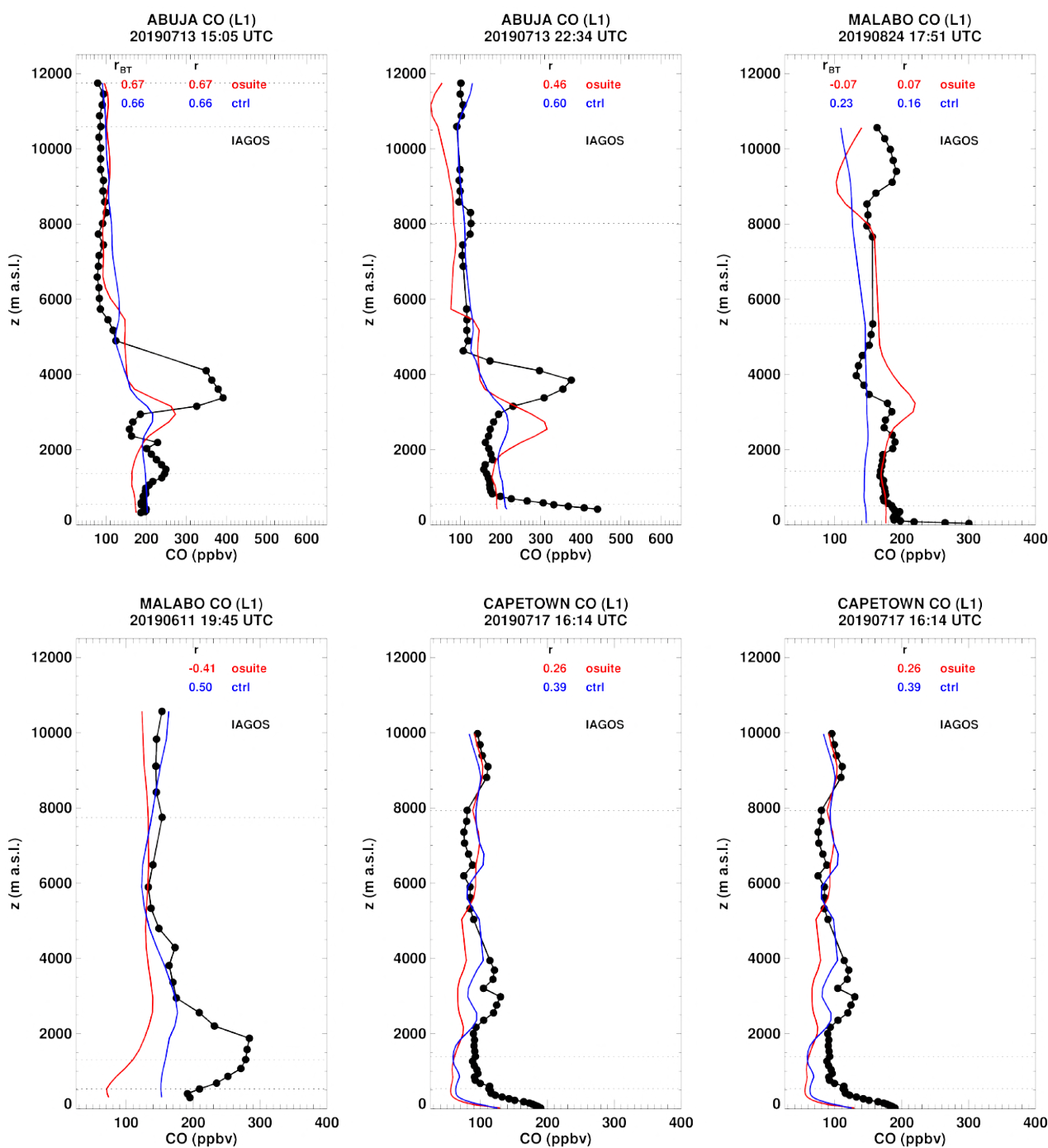
Figure 4.2.4a. Daily profile for CO from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over the Middle East during JJA 2019.
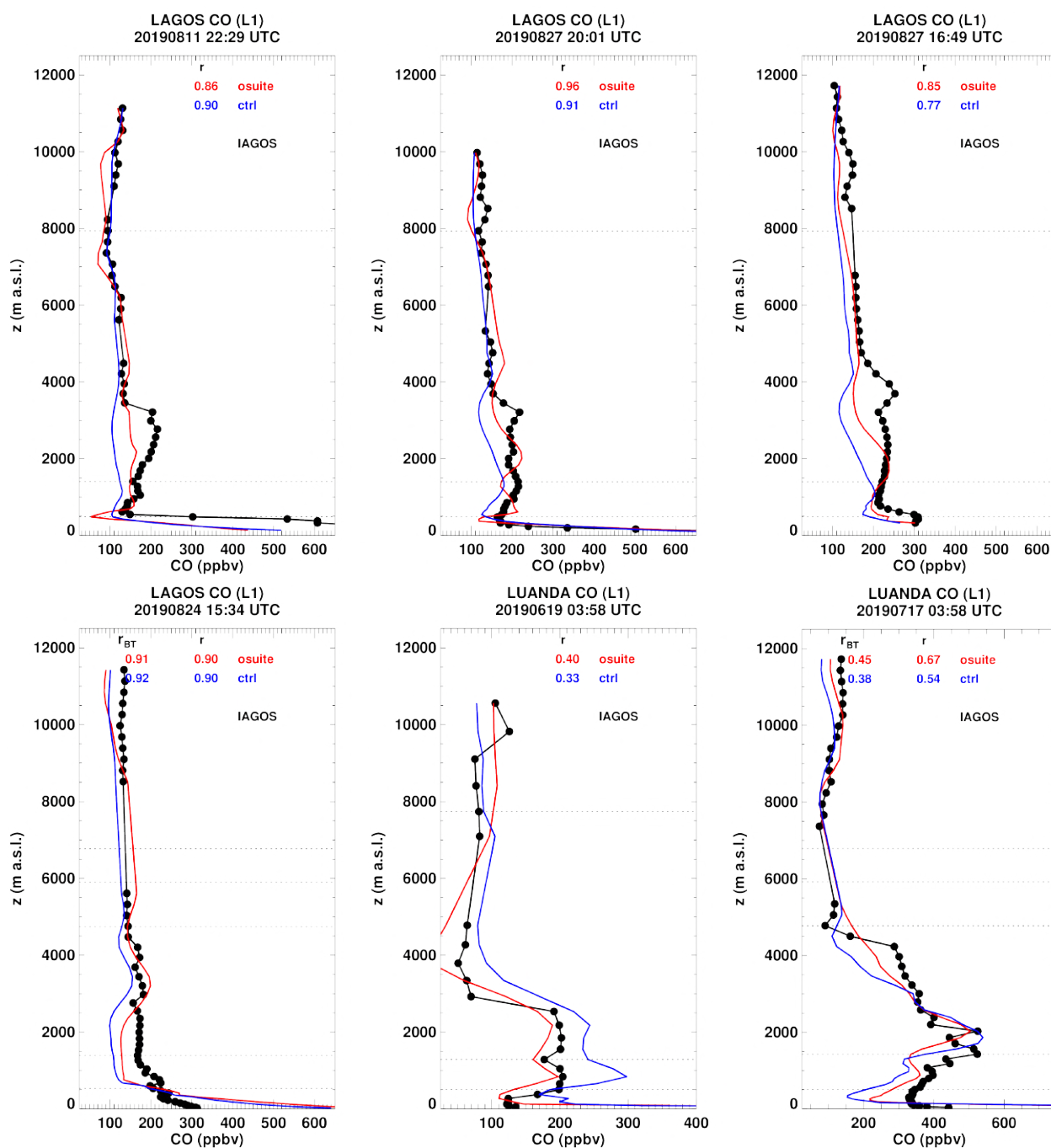
Figure 4.2.4b. Daily profile for CO from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over the Middle East during JJA 2019.
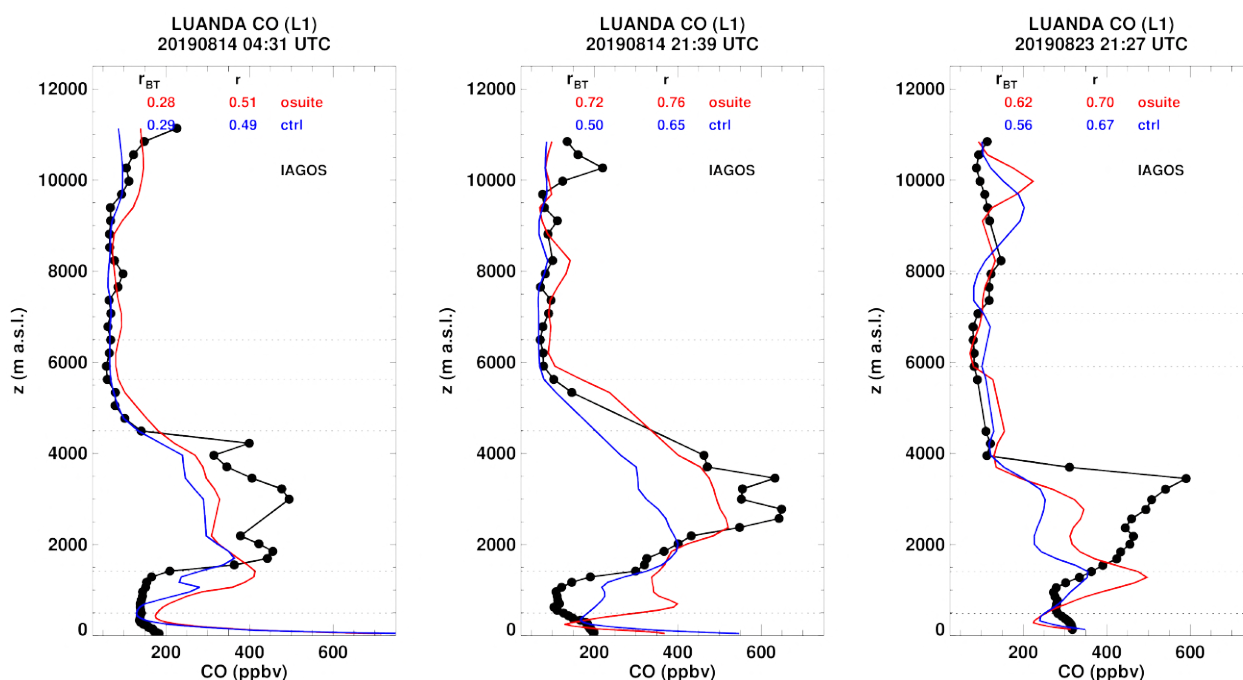
Africa

During JJA 2019, CO profiles are available over different regions of Africa (Fig. 4.2.5a-c) over Western Africa (Lagos, Malabo, Abuja, Luanda) but also over Eastern Africa (Nairobi) and South Africa (Cape Town). CO values are mostly underestimated from the surface to the free troposphere by both models, although surface values are sometimes overestimated at some airports (Lagos, Luanda). High mixing ratios of CO near the surface (> 400 ppbv) are in general well reproduced by the models (Lagos 11 and 27 August, Luanda 17 July). The plumes of CO from forest fires are well detected by the models, and the complex shape of the related profiles is generally well reproduced. However, the altitudes of the maxima in the free troposphere at Abuja are found at a slightly lower altitude by the models than observations. At all airports the performance of the o-suite is better than that of control run both regarding shape and magnitude, except for the southern hemisphere airport of Cape Town where control run present slightly better results than the o-suite. In the UTLS, o-suite and control run behave similarly at all locations and overall agree well with observations.

Figure 4.2.5a. Daily profile for CO from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over Western Africa during JJA 2019.

Figure 4.2.5b. Daily profile for CO from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over Western Africa during JJA 2019.

Figure 4.2.5c. Daily profile for CO from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over Western Africa during JJA 2019.

North America

CO profiles are available at many different locations over North America (Fig. 4.2.6a-c). In the profiles presented here, surface values of CO range between 150 ppbv to nearly 400 ppbv. For most profiles, surface and boundary layer values are underestimated by the two runs which often present similar performances. High mixing ratios of CO at the surface are largely underestimated by the two models as shown at the airport of Atlanta on 5 June and San Jose on 19 July.

During this summer, the unusually intense and long-lived wildfires that have taken place in the Arctic Circle have been monitored by CAMS (https://atmosphere.copernicus.eu/cams-monitors-unprecedented-wildfires-arctic) and most of them burned in Alaska and Siberia. At the beginning of July, the transport of pollution from wildfires in Alaska and Northern Ontario have affected several areas in Canada, as well as the mid-western and north-eastern states of the US such as Massachusetts. In the very last days of July the smoke from the fires in Siberia had spread into Alaska and started drifting east towards Canada and the US (https://www.nasa.gov/image-feature/goddard/2019/siberian-smoke-heading-towards-us-and-canada). The transport of plumes from arctic wildfires is observed by IAGOS on several dates at the airport of Boston, on 14-15 July, on 3 and 10 August. The satellite images available at https://worldview.earthdata.nasa.gov clearly show the arctic origin of these CO plumes. For most of these cases, both model runs detect these plumes but have different performances. On 14-15 July, a maximum of CO is observed near 4000-4500 m on several profiles reaching 350 ppbv at 2:29 UTC. The o-suite performs well for the first two profiles (14/07 UTC, 15/07 2:29 UTC) and largely underestimates the plume magnitude for the third profile on 15/07 at 23:26. The control run is modelling a very weak increase in CO for all these three profiles. For the event of the 3rd August, the shape of the profile present two maxima in the
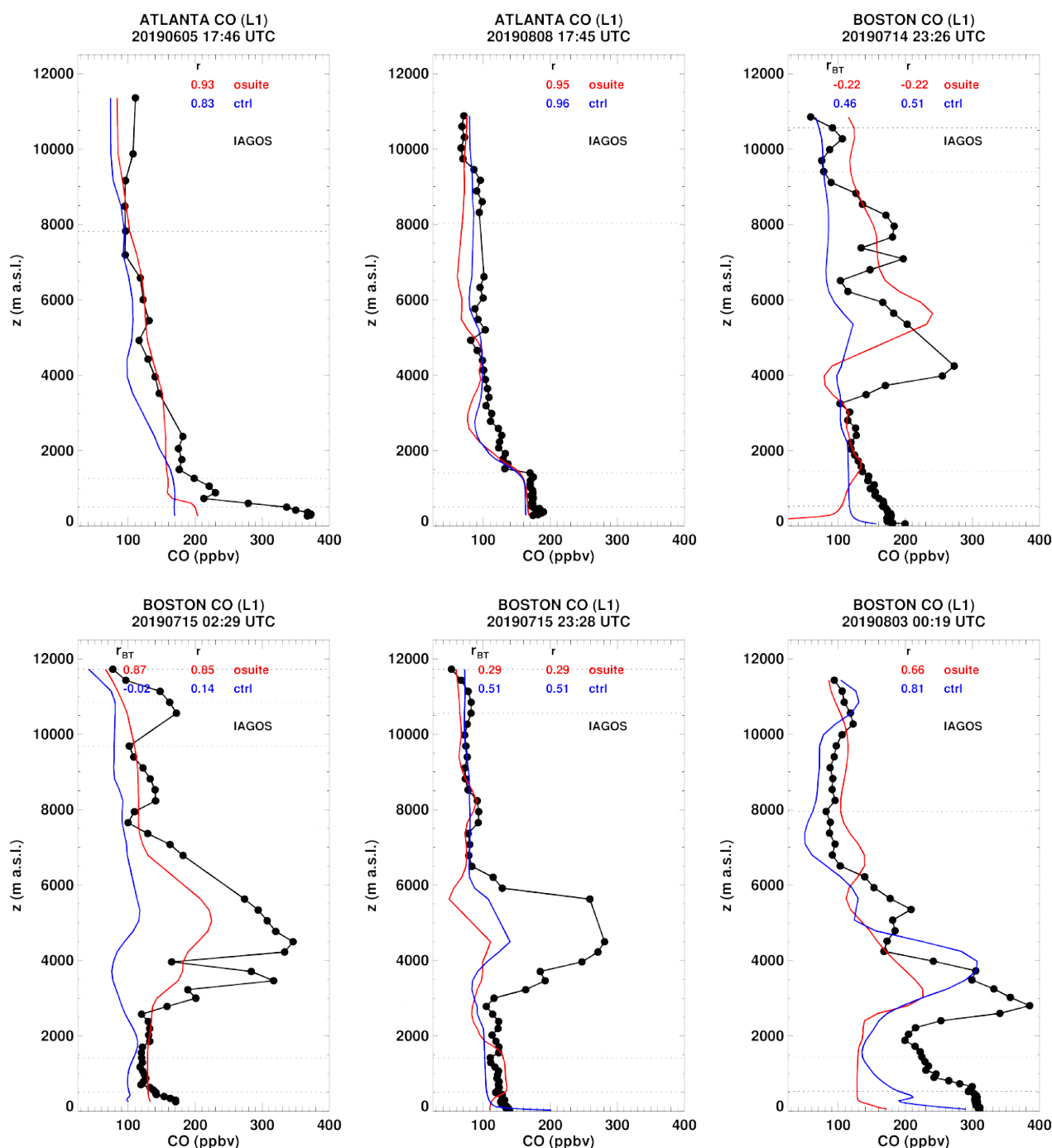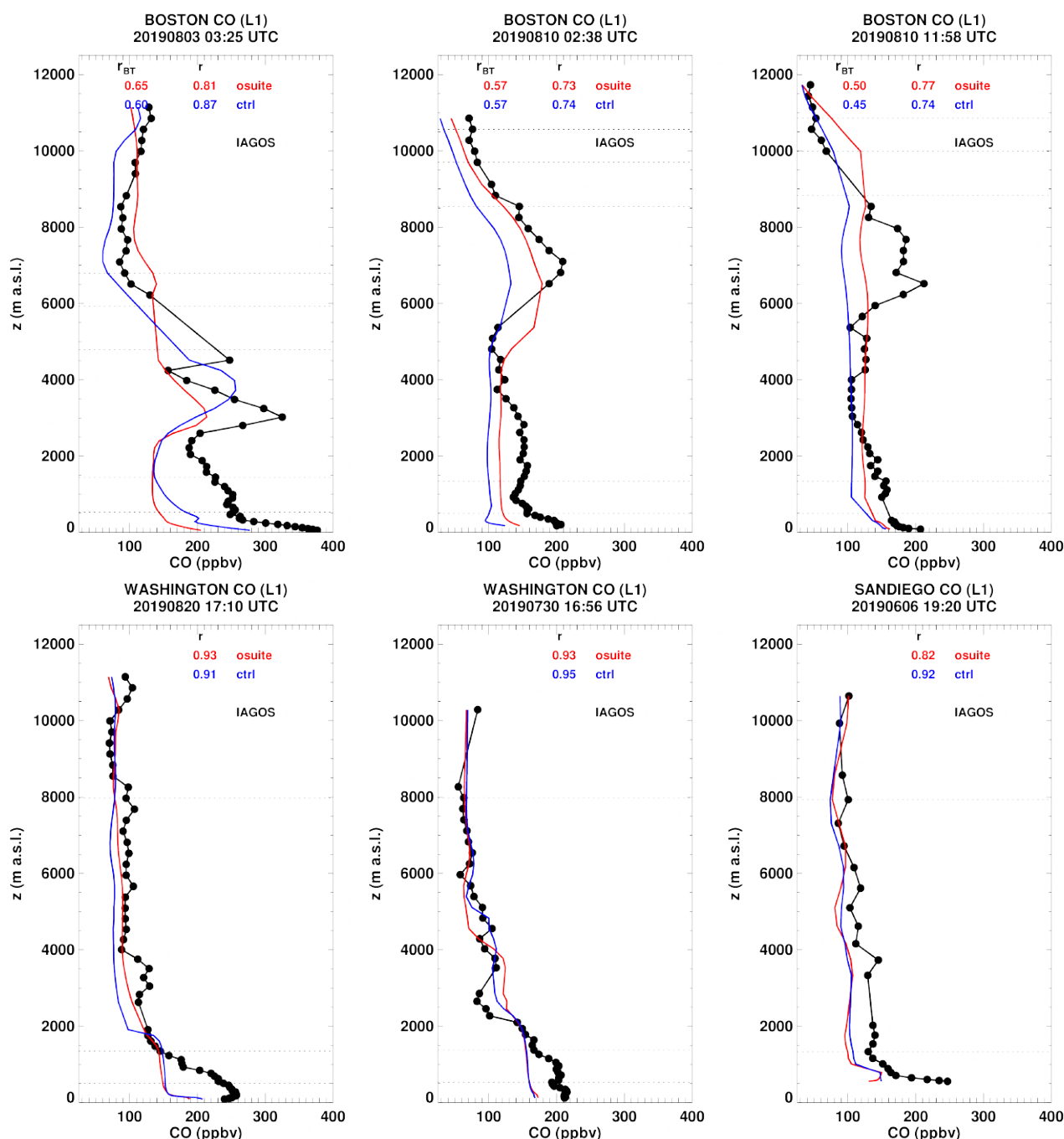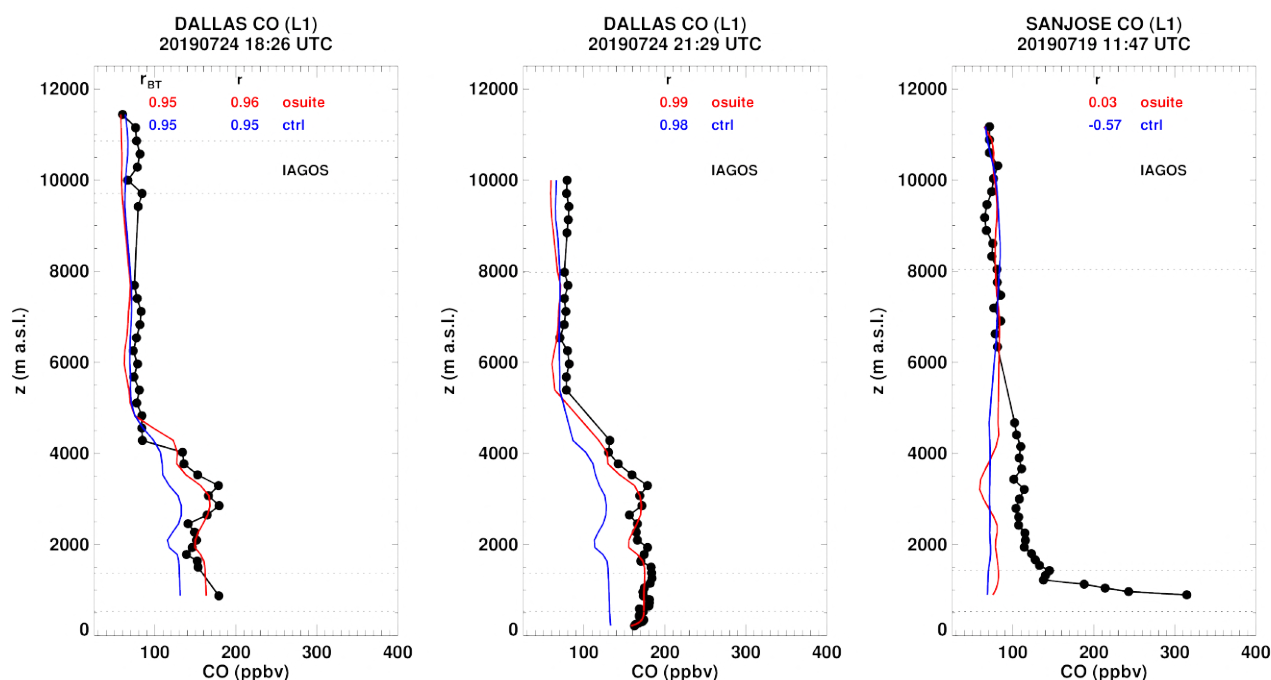
Figure 4.2.6a. Daily profile for CO from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over North America during JJA 2019.

free troposphere, the main one is observed near 3000 m with 400 ppbv at 00:29 UTC and 350 ppbv at 03:25. Both runs show well the plume with a slightly better agreement from control run, although the agreement is worse for the secondary maximum. In these two cases the altitude of the peaks is slightly higher for the models than for the observations. On 10 August the maximum is observed at higher altitudes (near 7 km) than for the two previously described events. At 2:38 UTC,

Figure 4.2.6b. Daily profile for CO from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over North America during JJA 2019.

the plume is detected at the correct altitude by both models with a better performance from the o-suite. For the other profile at 11:58 UTC, the plume is not seen in the control run while the o-suite shows a slight increase and underestimates the observed values. In the UTLS, o-suite and control run results are generally similar, showing a good agreement with observations except in the case of transport of pollution from forest fire plumes as it is the case in the profiles of Boston where the results differ more between the two runs with sometimes overestimations by the o-suite (14/07 23:26, 10/08 11:58).
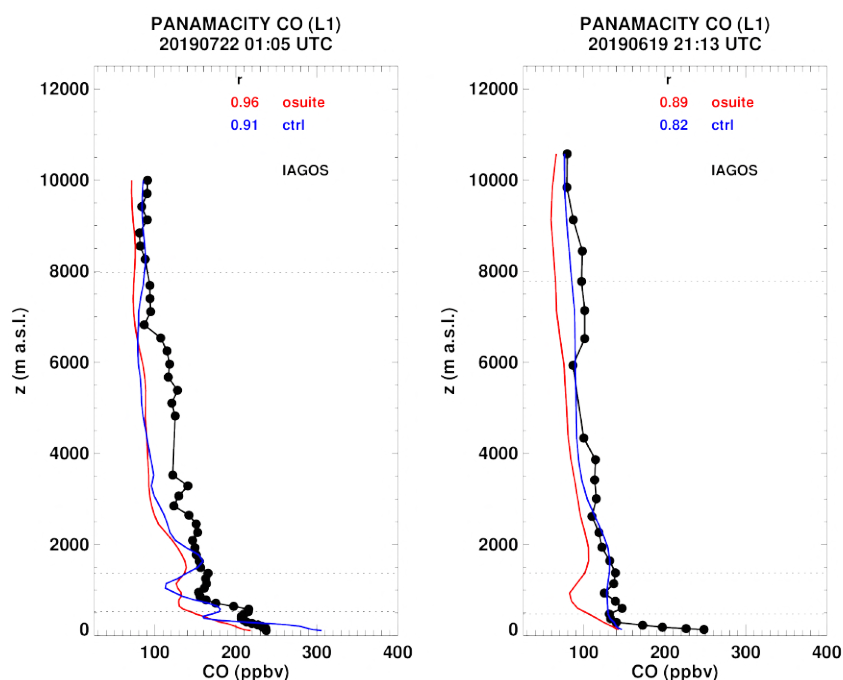
Figure 4.2.6c. Daily profile for CO from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over North America during JJA 2019.
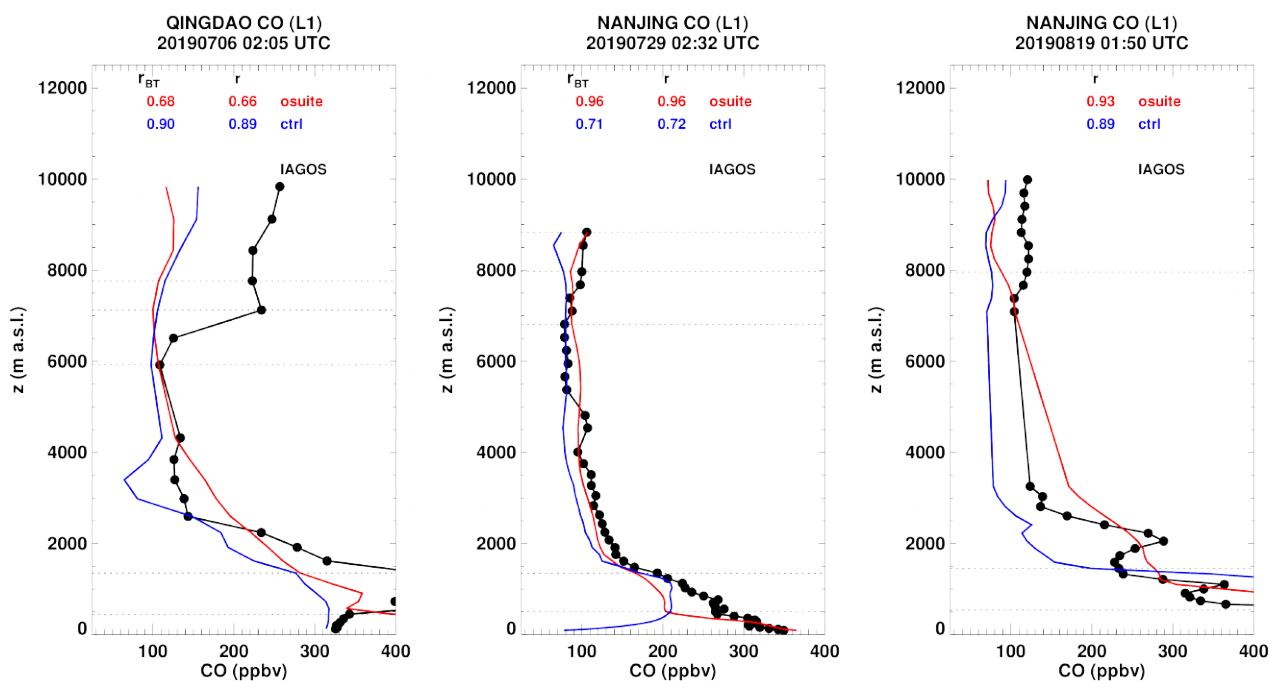
Figure 4.2.7a. Daily profile for CO from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over East Asia during JJA 2019.
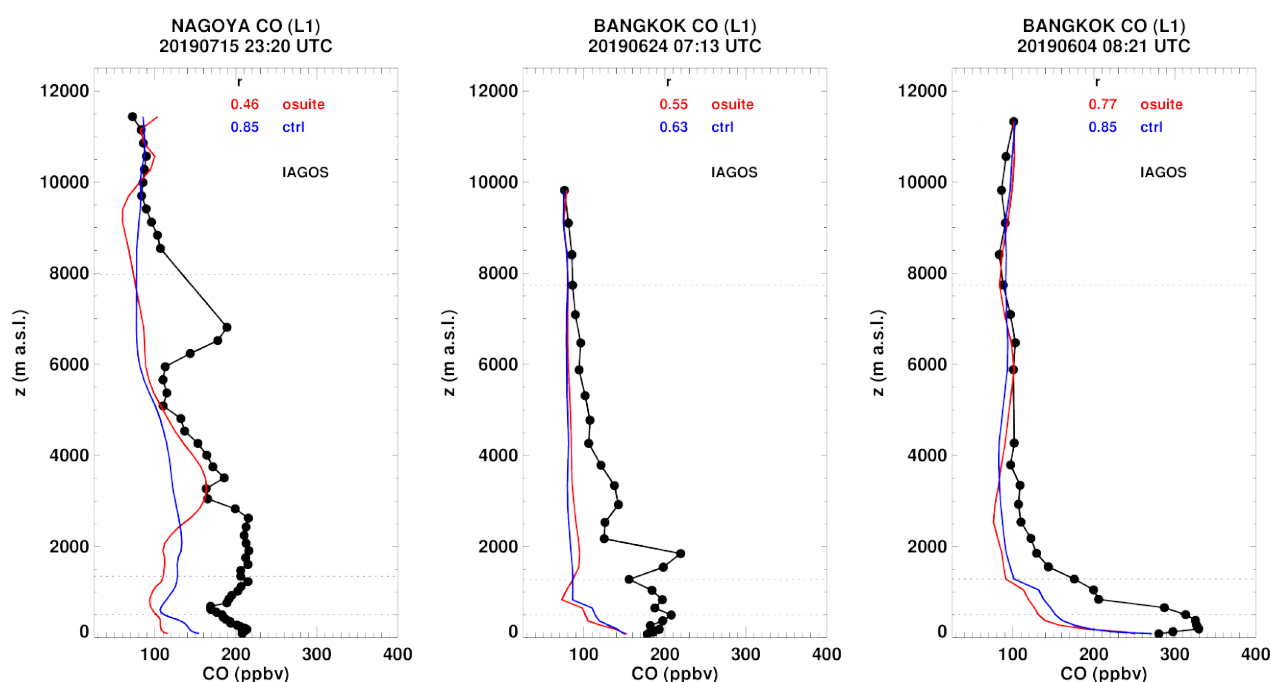
Figure 4.2.7b. Daily profile for CO from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over East Asia during JJA 2019.

Eastern Asia

CO profiles at the Asian airports of Qingdao, Nanjing, Nagoya, and Bangkok are shown in Fig. 4.2.7a-b. Over the Chinese airports (Qingdao and Nanjing) high CO mixing ratios are often observed near the surface and in the boundary layer, and the results from the o-suite are generally better than those from the control run, both concerning shape and magnitude (largest underestimation by the control run). At the airports of Nagoya (Japan) and Bangkok (Thailand), CO is mostly underestimated from the surface to the mid-troposphere. In the UTLS the performance of the models is similar at all locations except for the profile at Qingdao where high CO values reach 250 ppbv.

Central America

During JJA-2019, a few profiles of CO are available at Panama City, as shown in Fig. 4.2.8. The available profiles for this period show similar shapes with CO values of about 200 ppbv in the surface layer, and values in the range of 100-150 ppbv in the boundary layer. For the profiles at this airport the results from both runs are very similar in general over the period with slightly better performance by the control run in the low troposphere. In the UTLS, CO values are of about 100 ppbv for all profiles and both runs agree well with observations.

Figure 4.2.8. Daily profile for CO from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over South America during JJA 2019.

4.3 Validation against FTIR observations from the NDACC network

In this section, we compare the CO profiles of the CAMS products with FTIR measurements at 21 FTIR stations within the NDACC network. These ground-based, remote-sensing instruments are sensitive to the CO abundance in the troposphere and lower stratosphere, i.e. between the surface and up to 20 km altitude. Tropospheric and stratospheric CO partial columns are validated. A description of the instruments and applied methodologies can be found at http://nors.aeronomie.be.

Figure 4.3.1 show that the o-suite tropospheric columns of CO agree well. The model upgrade (60 to 137 levels) implemented in July 2019 changes the overall biases in both the troposphere and stratosphere. The bias for the tropospheric columns becomes -4% in JJA (-2% in MAM) and becomes comparable to the measurement uncertainty. The stratospheric column bias reduces to +6% in JJA (+10% in MAM) and now falls within the uncertainty.

The control run changes sign going from the NH (underestimation) to the SH (overestimation).

Fig. 4.3.2 shows a trend in the tropospheric CO column at Jungfraujoch (4km – TP) of about 1.5% per year. A similar trend is observed at Zugspitze (3 km above sea level), but not at other non-mountain sites like St Petersburg. The trend at the o-suite 1dFC at both mountain stations is much lower (around -0.5%/y), which suggests the trend is located in the upper tropospheric column and is related to the assimilation.
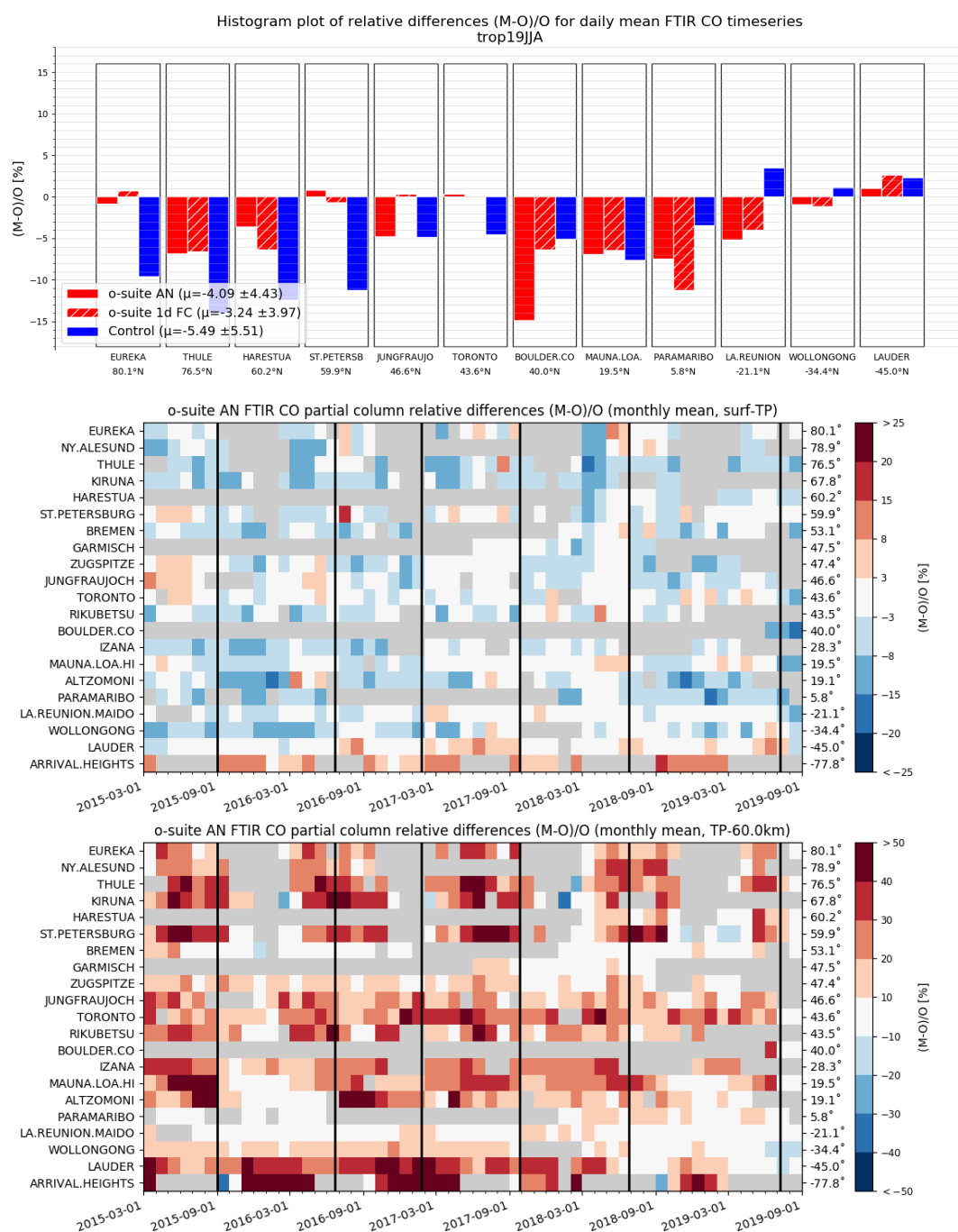
Figure 4.3.1: Seasonal relative mean bias for tropospheric CO columns (MB, %) for the considered period 2019 JJA (top) and monthly mean biases for a longer time period for the tropospheric CO columns (middle) and stratospheric CO columns (bottom) (model upgrades are indicated in black vertical lines).The overall uncertainty for the CO measurements is approximately 3% on the tropospheric columns and 10% for the stratospheric columns. The o-suite analysis averaged bias in tropospheric columns for all stations is -4% for JJA 2019. The bias in the stratosphere reduced to +6% and lies within the measurement's uncertainty. Stations are sorted with decreasing latitude (northern to southern hemisphere).
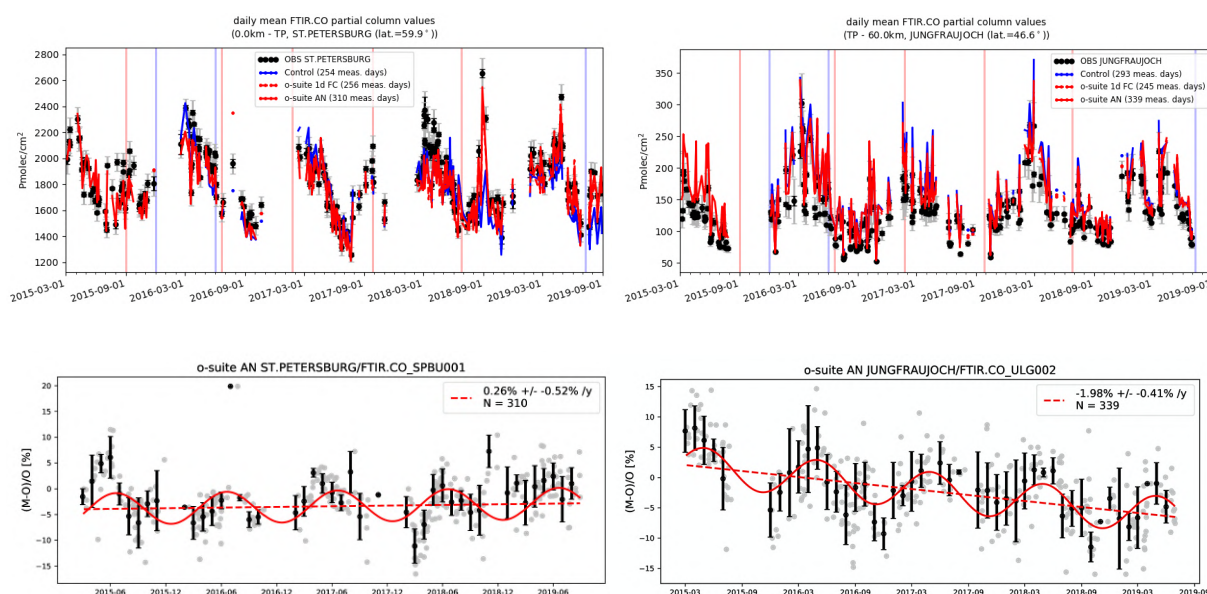
Figure 4.3.2: Top: daily mean values of tropospheric CO columns by the o-suite (AN and 1d FC, red) and the Control run (blue) compared to NDACC FTIR data at St Petersburg and Jungfraujoch for the period March 2015-August 2019. During March 2018 the o-suite underestimated the CO columns at St. Petersburg. Bottom row contains a linear fit and seasonal cycle fit through the relative differences for the o-suite analysis. An underestimation is observed during the local autumn/winter months. The negative trend at Jungfraujoch is not seen in the o-suite 1dFC.

The Taylor diagrams in Figure 4.3.3 provide information on the correlation of all three CAMS products under consideration with the FTIR time series. Leaving out the sites with few measurements, the assimilation has a positive effect on the correlation coefficient. Looking at the correlation values for the period 2019 JJA, the o-suite 1-day forecast (averaged correlation for all sites is 0.82) is comparable to the o-suite analysis (averaged correlation for all sites is 0.85).
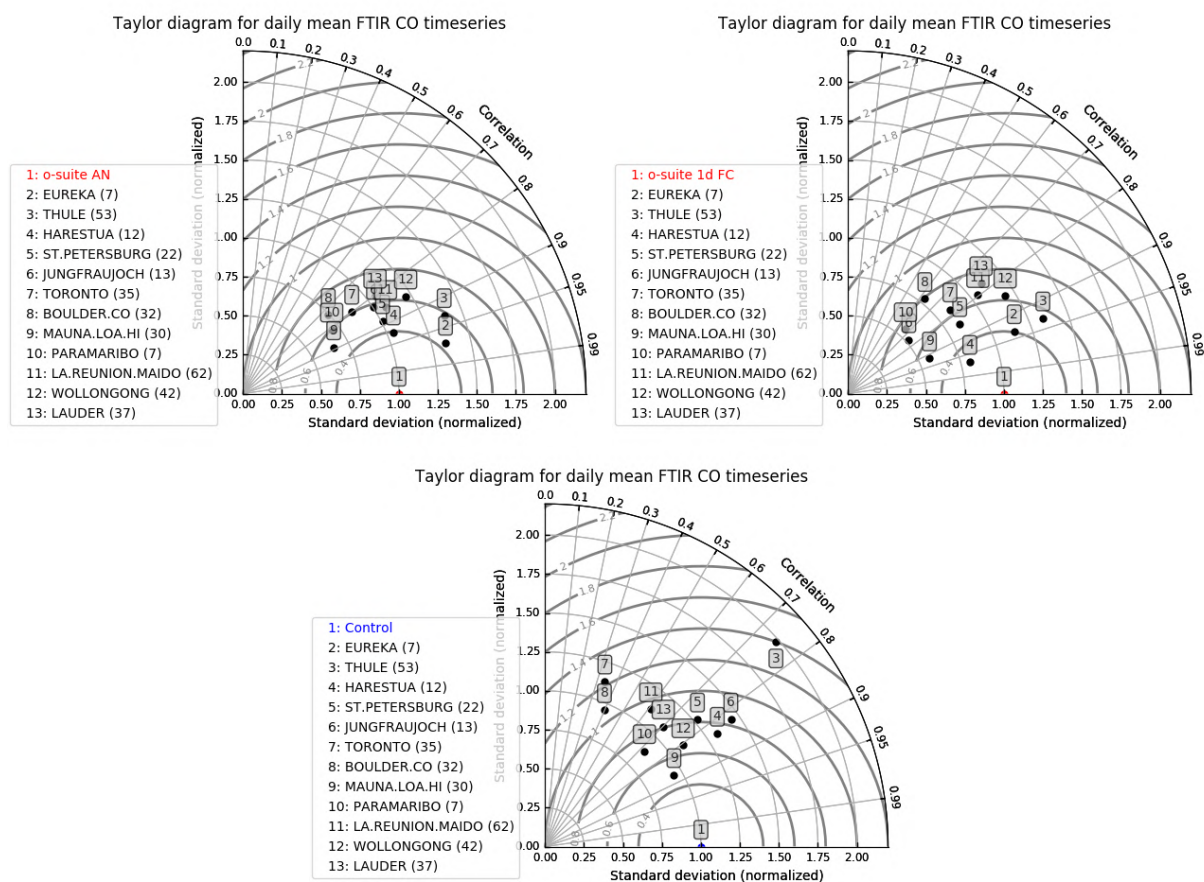
Figure 4.3.3: Taylor diagrams relating the standard deviations for the model /GB time series of tropospheric CO column data and their correlation. All timeseries are normalised such that the standard deviation of the model is 1. The variability of the CO columns in the o-suite 1-day forecast deviates more from the variability in the FTIR columns compared to the o-suite analysis.

4.4 Validation against FTIR observations from the TCCON network

CO column-averaged mole fractions of the CAMS models are compared with data from the Total Carbon Column Observing Network (TCCON). Column averaged mole fractions provide different information content than the in-situ measurements and are therefore complementary to the in-situ data.

In this section, we compare column averaged mole fractions of CO of the CAMS models with TCCON retrievals. Data from the following TCCON sites has been used: Izana (Blumenstock et al., 2017), Reunion (De Mazière et al., 2017), Bialystok (Deutscher et al., 2017), Manaus (Dubey et al., 2017), Four Corners (Dubey et al., 2017), Ascension (Feist et al., 2017), Anmeyondo (Goo et al., 2017), Darwin (Griffith et al., 2017), Wollongong (Griffith et al., 2017), Karlsruhe (Hase et al., 2017), Edwards (Iraci et al., 2017), Indianapolis (Iraci et al., 2017), Saga (Kawakami et al., 2017), Sodankyla (Kivi et al., 2017), Hefei (Liu et al., 2018), Tsukuba (Morino et al., 2017), Burgos (Morino et al., 2018), Rikubetsu (Morino et al., 2017), Bremen (Notholt et al., 2017), Spitsbergen (Notholt et al., 2017), Lauder (Sherlock et al., 2017, Pollard et al., 2019), Eureka (Strong et al., 2018), Garmisch

Figure 4.4.1: Monthly differences for the last 4 years. The stations are sorted by latitude (northern to southern hemisphere).
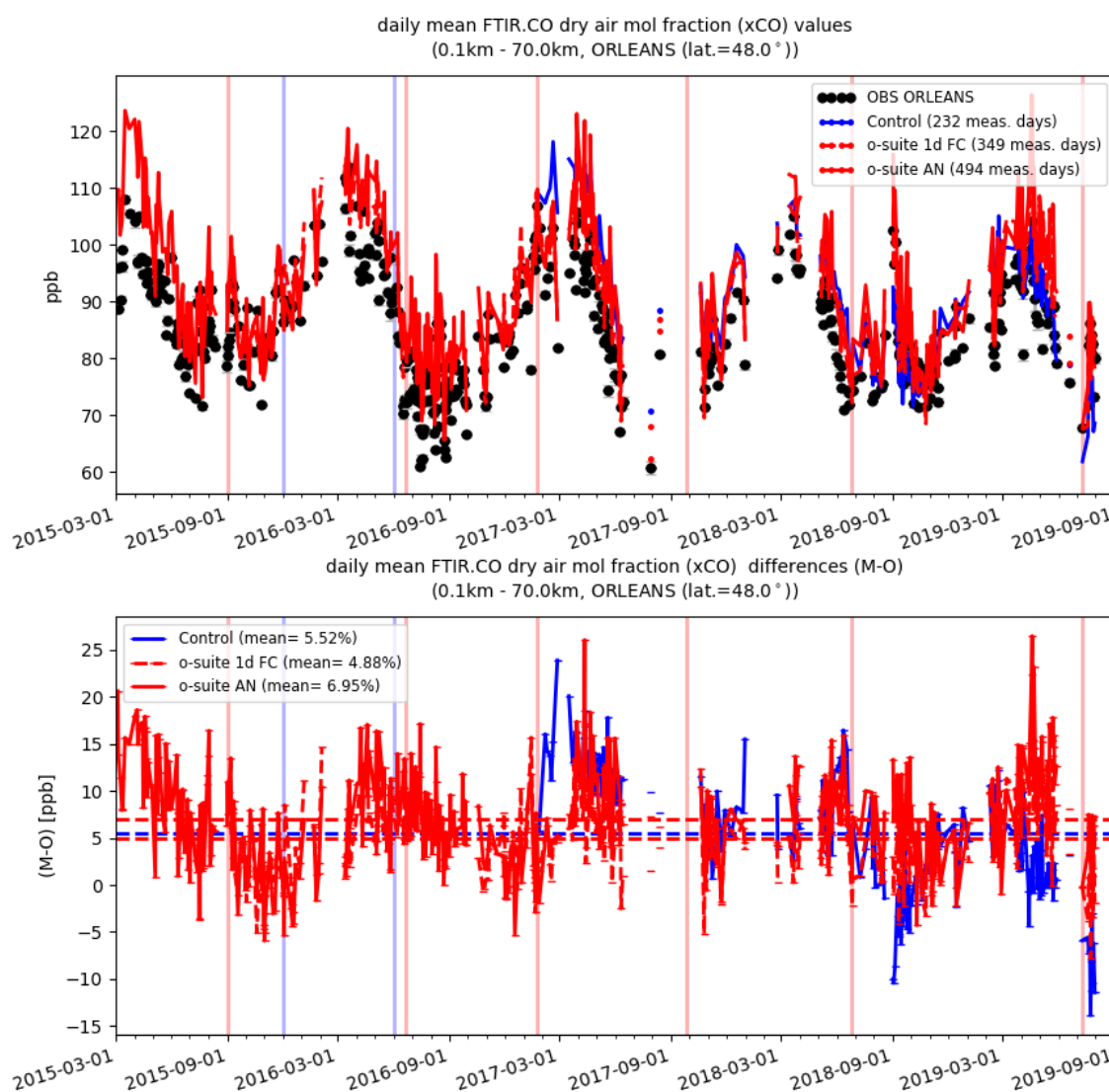
Figure 4.4.2: Comparison of the CO model data with TCCON CO at Orleans.

(Sussmann et al., 2017), Zugspitze (Sussmann et al., 2018), Paris (Te et al., 2017), Orleans (Warneke et al., 2017), Park Falls (Wennberg et al., 2017), Caltech (Wennberg et al., 2017), Lamont (Wennberg et al., 2017), Jet Propulsion Laboratory (Wennberg et al., 2017), East Trout Lake (Wunch et al., 2017)

For the validation of the models in June, July and August the sites that made data available for more than just a few days were Eureka, East Trout Lake, Orleans and Izana. Since TCCON PIs usually process the data in batches and the requirement within TCCON is to make the data publically available 1 year after the measurement, the availability of data for these reports is limited. Previously data from Bialystok, Orleans and Reunion was timely available for the validation of the CAMS models. The Bialystok site has stopped operation and the instrument has been transported to Cyprus, where measurements were started in September 2019. These data will be available for the next report. It can therefore be expected that for future reports at least data from Cyprus, Orleans and Reunion will be available.
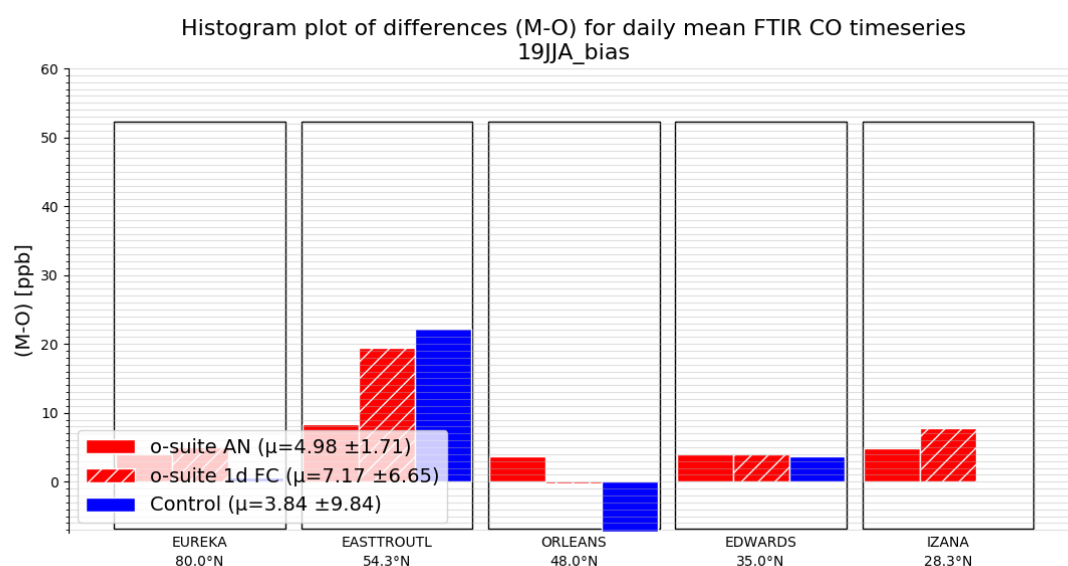
Figure 4.4.3: Differences during the reporting period. The different sites cover different periods of the comparison period. Therefore, only a qualitative comparison can be done.

The comparisons show that all models capture the seasonality well and the agreement is within 5%. For Orleans the comparison is shown in Fig 4.4.2. While there is an overestimation by all models for most of the comparison period, the control run shows an underestimation during the last days of the comparison. During this period a strong increase of the measured CO is observed, which is most likely due to transported CO from biomass burning.

4.5 Evaluation with MOPITT and IASI data

In this section, modelled CO total columns are compared to MOPITT version 8 (thermal infrared radiances) (Emmons et. al., 2009, Deeter et al., 2010) and IASI satellite retrievals (Clerbaux et al., 2009). Figure 4.5.1 shows the global distribution of CO total columns retrieved from MOPITT V8 (top left) and IASI (top right) and the relative biases of the model runs with respect to MOPITT V8 for July 2019.

The difference between MOPITT and IASI data can be partly explained by the differing averaging time period due to the data availability. MOPITT data are not available for the last week of July and IASI data are not available for the first week of the month. MOPITT shows high values over the biomass burning area in Central Africa and over the north-eastern part of China, and enhanced values over Siberia and Canada, where IASI shows much higher values.

The modelled CO geographical distribution and magnitude of values shows that the model performs reasonably well. The relative difference between the model runs and MOPITT shows that the o-suite performs better than the control run without data assimilation. The bias of the o-suite run is within 10% with some regional exceptions where the negative bias reaches 20% (mostly over land). The control run shows an underestimation of the satellite data over the Northern Hemisphere continents up to 30% and overestimations over the tropics. Both runs show a growing positive bias
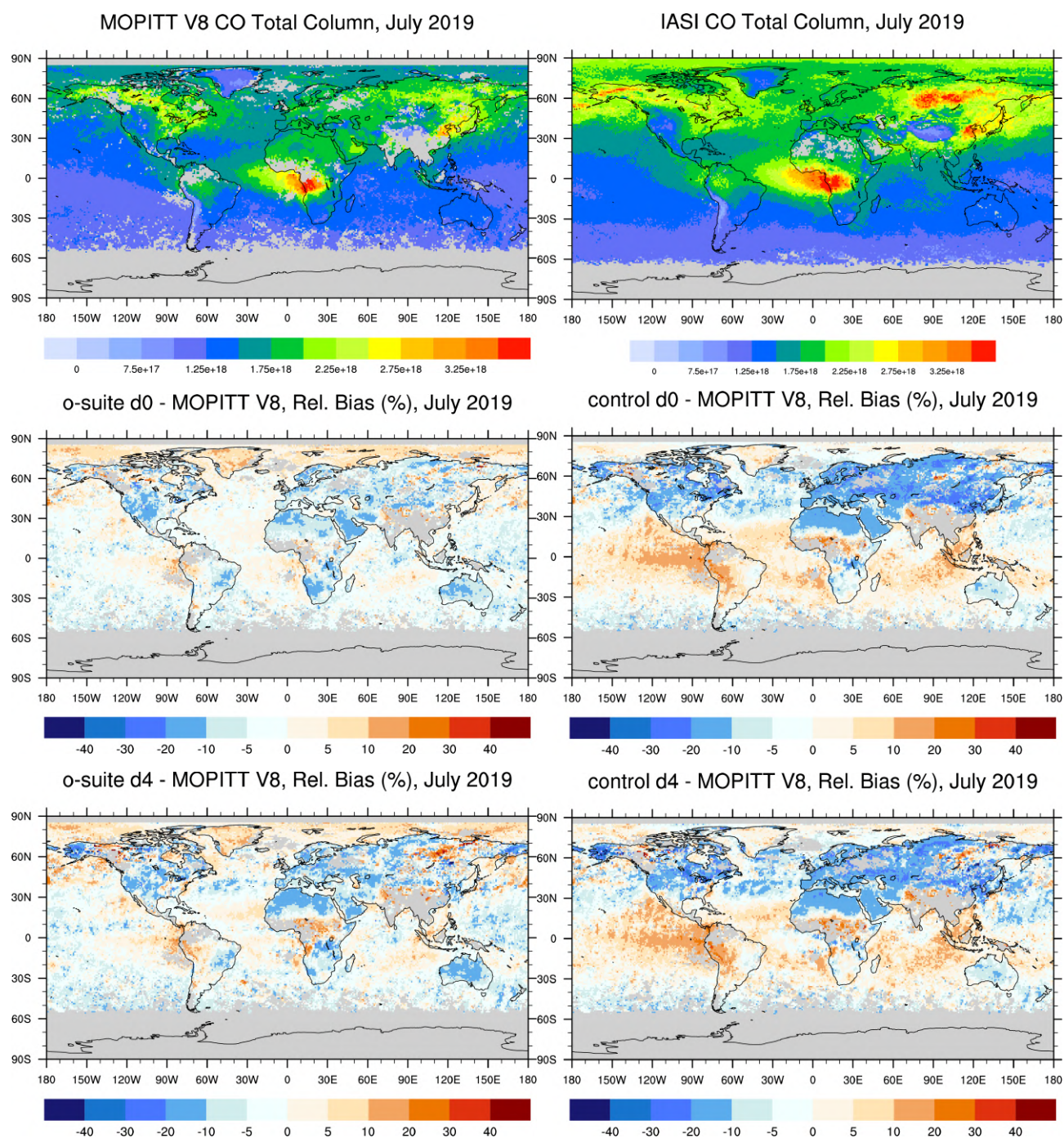
Fig. 4.5.1: CO total columns for MOPITT V8 (top left) and IASI (top right) satellite retrievals and relative difference between the model runs and MOPITT for April 2019: o-suite (middle left), control run (middle right), o-suite 4th forecast day (bottom left), o-suite 4th forecast day (bottom right). Grey colour indicates missing values.

on the 4th forecast day over the fire active areas in Siberia, Canada and Central Africa, as well as a growing negative bias in some other areas over the continents.

Figure 4.5.2 shows time series of CO total column for MOPITT V8, IASI and the model runs over the eight selected regions. For the comparison with MOPITT, the modelled CO concentrations were transformed using MOPITT V8 averaging kernels (Deeter, 2004). Both, MOPITT and IASI CO total columns are assimilated in the o-suite run, while a bias correction scheme is applied to IASI data to bring it in line with the analysis. MOPITT and IASI CO total columns show a relatively similar variability over different regions. IASI CO values are lower than MOPITT over most regions with some seasonal exceptions till the year 2016. Since then IASI and MOPITT are more consistent with each other over Europe, the US and East Asia. Significant difference between MOPITT and IASI are observed over the Alaskan and Siberian fire regions in winter seasons, with IASI CO total column values being lower up to 30%. In North and South Africa, deviations become larger since 2016 with IASI values being higher than MOPITT by up to 20%. The modelled seasonality of CO total columns is in relatively good agreement with the retrievals. In general, the comparison between the o-suite and control run shows that the assimilation of satellite CO has a more positive, pronounced impact on model results over East and South Asia, South Africa, and since the end of 2016, over the US in winter and spring seasons, and smaller impact over the other regions. Since June 2016, the o-suite shows very good agreement with the satellite retrievals over Europe and the US with biases less than 5%. In late summer and early autumn of 2018 over Europe, the control run has larger negative biases compared to the satellite data then early in 2018 and the two previous autumn seasons.

A general reduction of CO values from 2015 to 2018 can be seen over Europe, the US and East Asian regions. The South African region shows a slight increase of the seasonal minimum compared to previous springs.

The modified normalized mean bias (MNMB) of the model runs compared to MOPITT V8 (Fig. 4.5.3) allows quantifying the impact of the assimilation on the model performance. The o-suite model run shows negative biases over Europe, the US and Alaskan fire regions with some seasonal exceptions.

The control run shows a systematic positive bias up to 20% over South Asia in November-December 2014, 2015, 2016, and 2017. Over southern Africa the control run overestimates satellite retrieved values by up to 25% in winter and spring 2015, 2016, and 2017. In general, the o-suite is within +/- 10% in all regions, while the control run shows larger biases over East and South Asia and North and South Africa, as well as stronger seasonal cycles.

In Europe and US regions, the o-suite performs similar to the previous year showing negative biases within 5-10%, while the negative bias of the control run is slightly increased. In East Asia, the o-suite bias is within 2% and the control run's bias increases to -10%. In South Asia, the o-suite bias changes sign to negative (but stays within 5%) and the control run's data improve and are very similar to the observations. Siberian and Alaskan fire regions are negatively biased for both, the o-suite (within 5%) and the control runs (10-15%). In South and North Africa, the o-suite is negatively biased (within 10%).
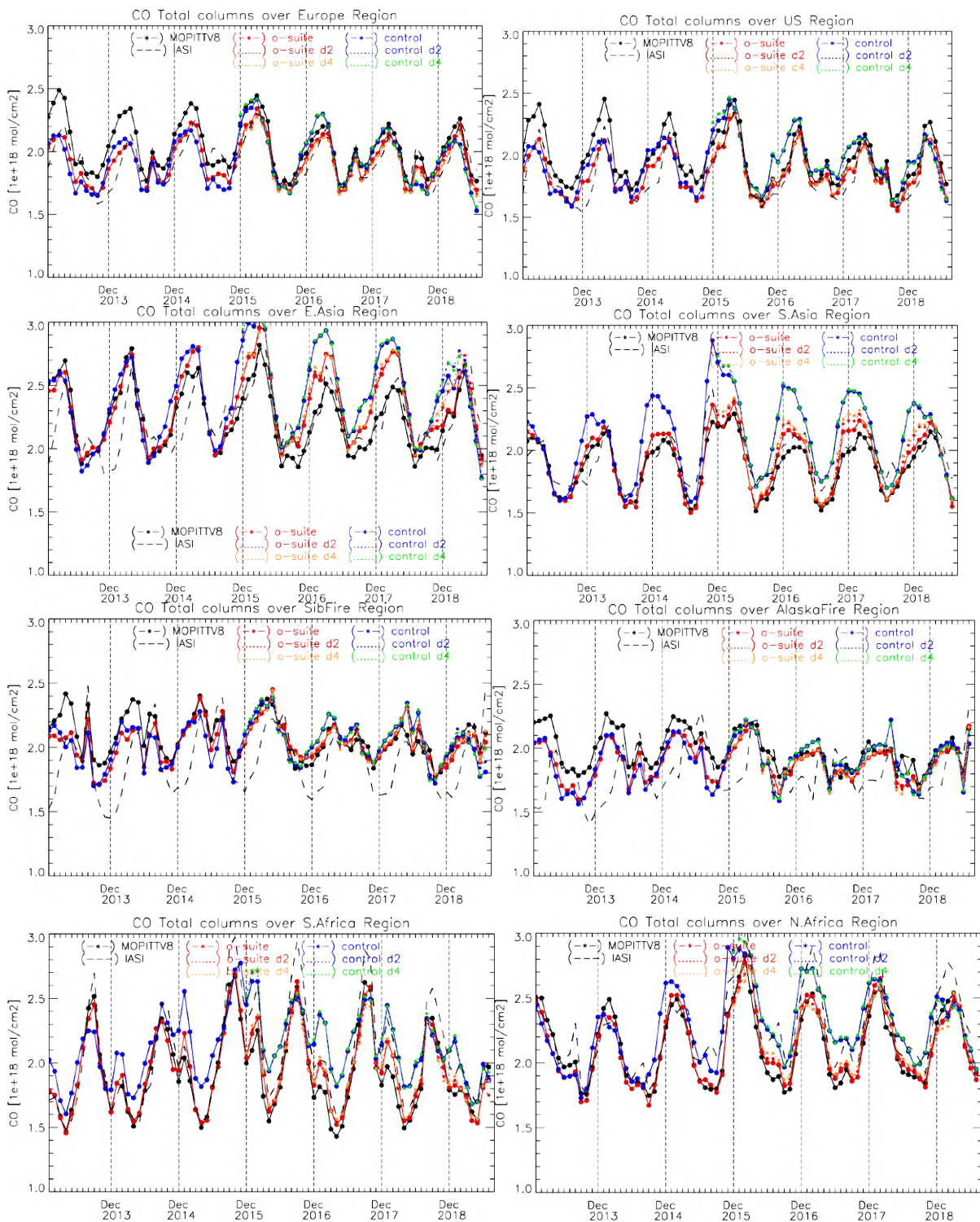
Fig. 4.5.2: Time series of CO total columns for satellite retrievals MOPITT V8, IASI (black) and the model runs over the selected regions: o-suite (red, solid), control (blue, solid), o-suite 2nd forecast day (red, dotted), o-suite 4th forecast day (orange, dotted), control 2nd forecast day (blue, dotted), control 4th forecast day (green, dotted). Period: January 2013 to July 2019.
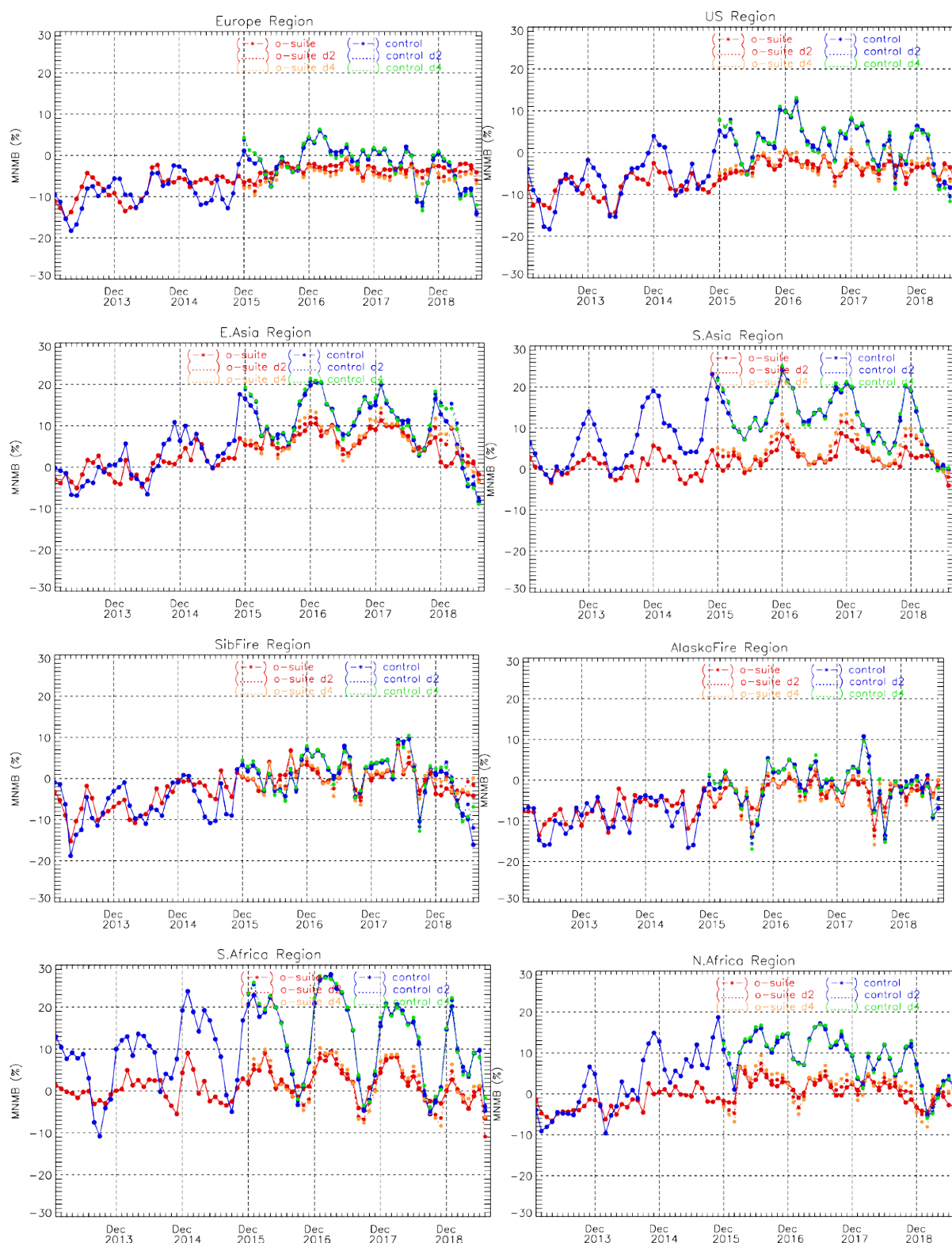
Fig. 4.5.3: Timeseries of modified normalized mean bias (%) for CO total columns from the model simulations vs MOPITT V8 retrievals over selected regions. O-suite (red, solid), control run (blue, solid), o-suite 2nd forecast day (red, dotted), o-suite 4th forecast day (orange, dotted), control 2nd forecast day (blue, dotted), control 4th forecast day (green, dotted). Period: January 2013 to July 2019.

5. Tropospheric nitrogen dioxide

5.1 Evaluation against GOME-2 and TROPOMI retrievals

In this section, model columns of tropospheric NO_2 are compared to SCIAMACHY/Envisat NO_2 satellite retrievals (IUP-UB v0.7) [Richter et al., 2005] for model data before April 2012, and to GOME-2/MetOp-A NO_2 satellite retrievals (IUP-UB v1.0) [Richter et al., 2011] for more recent simulations. First comparisons to TROPOMI/Sentinel-5P data (IUP-UB v0.1, preliminary) are provided, using the CAMS o-suite as a-priori in the TROPOMI retrievals. This satellite data provides excellent coverage in space and time and very good statistics. However, only integrated tropospheric columns are available, and the satellite data is always taken at the same local time, roughly 09:30 LT for GOME-2, 10:00 LT for SCIAMACHY and 13:30 LT for TROPOMI and at clear sky only. Therefore, model data are vertically integrated, interpolated in time and then sampled to match the satellite data. The satellite data were gridded to model resolution (currently 0.4° x 0.4° degree). Model data were treated with the same reference sector subtraction approach as the satellite data for all SCIAMACHY/GOME-2 comparisons. For all comparisons to TROPOMI satellite data, tropospheric NO_2 columns over the clean Pacific reference sector simulated by CAMS-global were added to the TROPOMI data. For TROPOMI comparisons before July 2019, the stratospheric contribution has been removed from the measurements according to the method by Hilboll et al. (2013) using simulations from the B3D-CTM (Sinnhuber at al., 2003a; Sinnhuber et al., 2003b; Winkler et al., 2008) scaled to satellite values over the clean Pacific reference sector. For July and August 2019, the reference sector method has been applied to the TROPOMI data. Uncertainties in NO_2 satellite retrievals are large and depend on the region and season. Winter values in mid and high latitudes are usually associated with larger error margins. As a rough estimate, systematic uncertainties in regions with significant pollution are on the order of 20% – 30%.

Figure 5.1.1 shows global maps of GOME-2 and model monthly mean tropospheric NO_2 columns as well as differences between retrievals and simulations for June 2019 as an example of the maps for summer 2019. The overall spatial distribution and magnitude of tropospheric NO_2 is well reproduced by both CAMS runs, indicating that emission patterns and NO_x photochemistry are reasonably represented. Some differences are apparent between observations and simulations, with generally larger shipping signals simulated by the models. For example, shipping signals are much more pronounced in model simulations to the south of India. Emissions over Europe and especially the pollution hotspots around the Benelux countries are regularly underestimated, especially during winter. However, other local maxima of values observed over anthropogenic emission hotspots in East Asia (e.g. over the heavily populated Sichuan Basin; 30°N, 105°E), India and others such as Teheran, Mecca, around Lebanon/Israel and Moscow are regularly overestimated. Values over boreal forest fires in Siberia, Alaska, Canada are overestimated. Similar conclusions arise from TROPOMI based map comparisons shown in Figure 5.1.2. However, the overestimation of values over India and some parts of Southern Africa does not show up when using TROPOMI as a reference. Differences in comparison results are in principal due to differences in observation time or differences in the retrieval products.
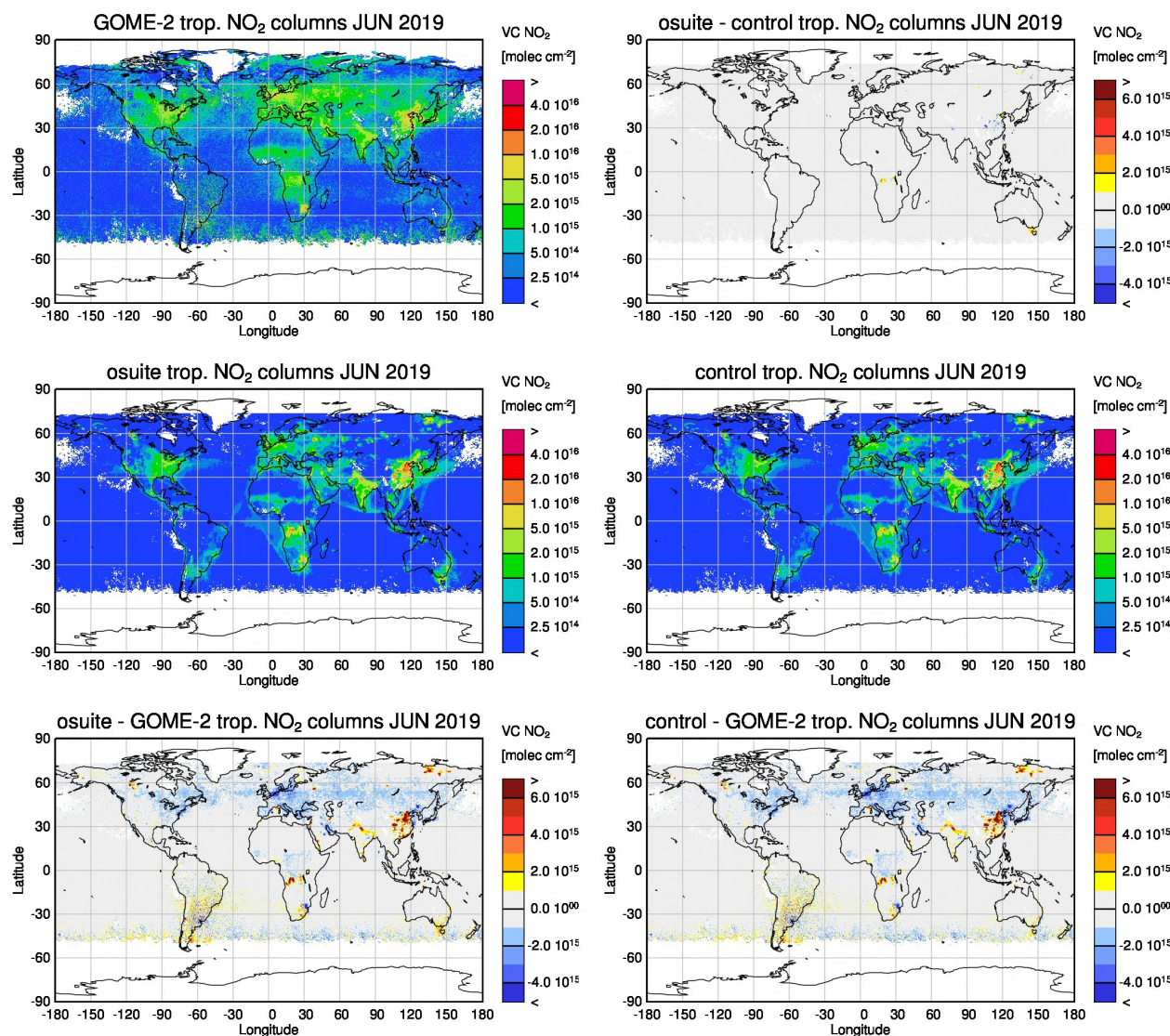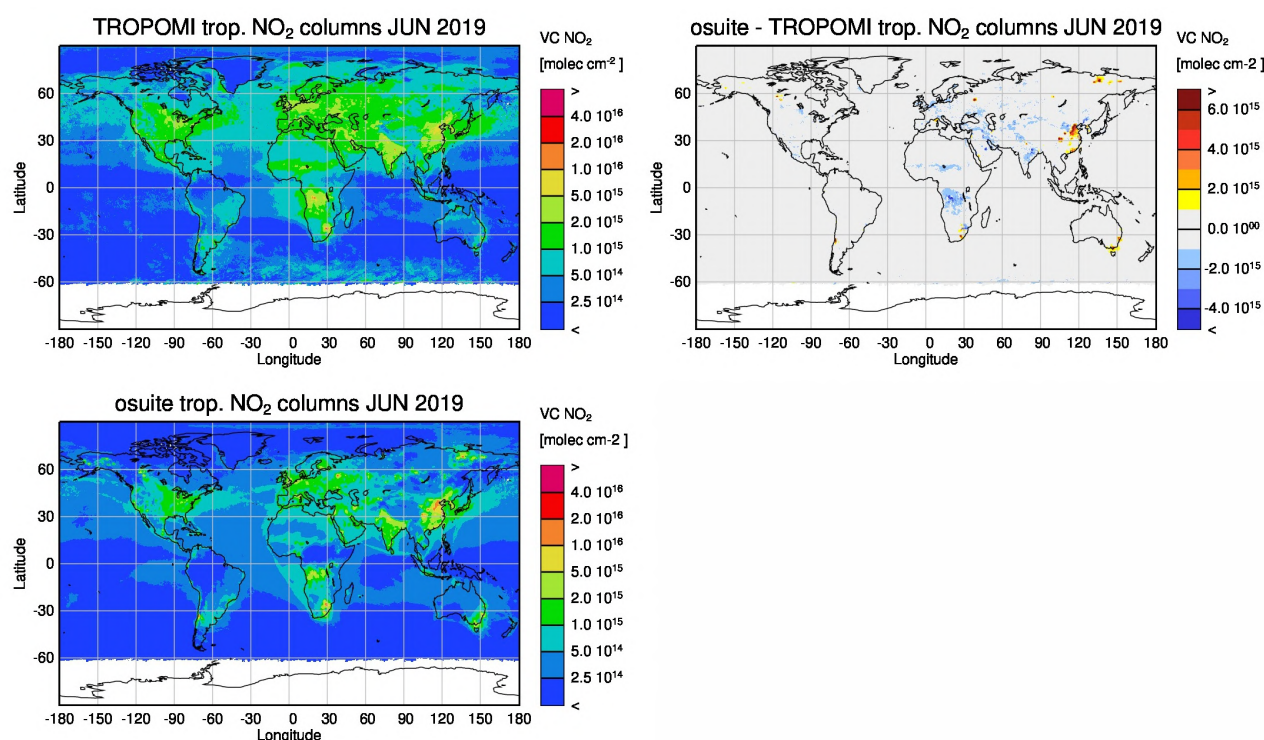
Figure 5.1.1: Global map comparisons of satellite retrieved, and model simulated tropospheric NO_2 columns [molec cm^{-2}] for June 2019. The top row shows monthly mean tropospheric NO_2 columns retrieved from GOME-2 as well as the difference between o-suite and control, the second row shows the corresponding tropospheric NO_2 columns for model simulated averages. The third row shows differences of monthly means between models and GOME-2. GOME-2 data were gridded to model resolution (i.e. 0.4° x 0.4° degree). Model data were treated using the same stratospheric correction method as for the satellite data.

Closer inspection of the seasonal variation of tropospheric NO_2 in some selected regions (Fig. 5.1.3) reveals significant differences between the models and points to some simulation problems. Over regions where anthropogenic emissions are major contributors to NO_x emissions, models catch the shape of the satellite time series rather well. However, over East-Asia absolute values and seasonality were strongly underestimated before 2014 by all model runs (most likely due to an underestimation of anthropogenic emissions) for all seasons apart from summertime minima, with the o-suite showing the best results since an upgrade in July 2012. As wintertime NO_2 column retrievals decreased significantly in 2014, model simulated wintertime maxima are in better agreement with the satellite retrieved ones for recent years. However, the observed NO_2 decrease

Figure 5.1.2: Global map comparisons of satellite retrieved, and model simulated tropospheric NO₂ columns [molec cm-2] for June 2019: (top left) TROPOMI, (top right) o-suite minus TROPOMI, (bottom left) o-suite. TROPOMI data were gridded to model resolution (i.e. 0.4° x 0.4° degree) and the CAMS o-suite was used as a-priori in the retrievals. Comparisons to the control are not available for this report.

is not reproduced by the simulations and therefore the better agreement for more recent years cannot be attributed to model improvements. Moreover, summertime model minima increased in 2015 compared to previous years, which is in contrast to the satellite retrievals, so that the simulated values for the summers since 2015 are about 50% larger than satellite retrieved ones.

As for East-Asia, a decrease in satellite retrieved values also occurred in 2015 over Europe where a peak is usually found around January, which was, as a result, only slightly underestimated by the models for January 2015. The underestimation of tropospheric NO₂ columns over Europe may be caused to some extent by a change of emission inventories in 2012. However, the situation changed for the three winter periods between 2015 and 2017, for which GOME-2 shows (compared to previous years) a strong increase in January peak values, combined with a decrease in values for December and February that is not reproduced by the models. It is not clear if the GOME-2 observations are realistic here, although an inspection of daily GOME-2 satellite images did not point to problems regarding the retrieval. The retrievals show the same pattern as the simulations however for winter 2018/2019.

Over regions where biomass burning is the major contributor to NOₓ emissions, seasonality and amplitude of model columns are determined by fire emissions. The seasonality for the two regions in Africa was simulated reasonably well for 2010 and after October 2011. In the time period in between, a bug in reading fire emissions lead to simulation errors for all MOZART runs. Over North-
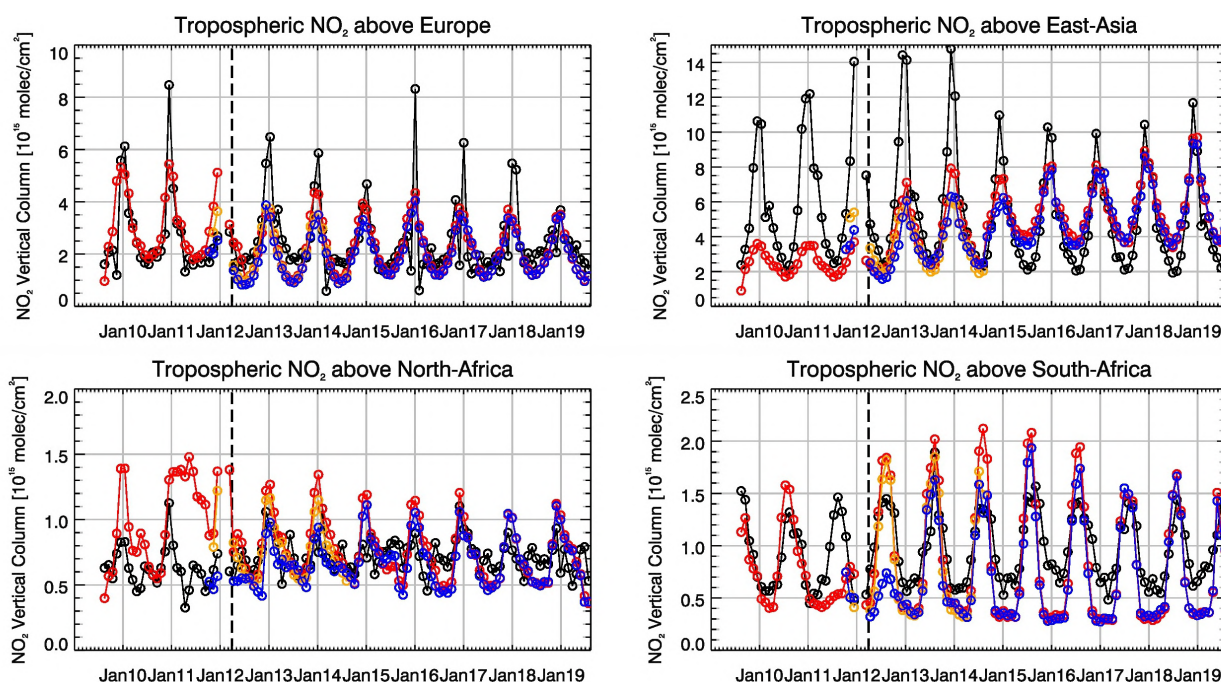
Figure 5.1.3: Time series of average tropospheric NO_2 columns [10^{15} molec cm^{-2}] from SCIAMACHY (up to March 2012, black) and GOME-2 (from April 2012 onwards, black) compared to model results (red: o-suite, blue: MACC_fncrt_TM5/MACC_CIFS_TM5/control, orange - MACC_fcnrt_MOZ) for different regions (see Annex 2 for definition of regions). The upper panels represent regions dominated by anthropogenic emissions, and the lower panels represent those dominated by biomass burning. Vertical dashed black lines mark the change from SCIAMACHY to GOME-2 based comparisons in April 2012.
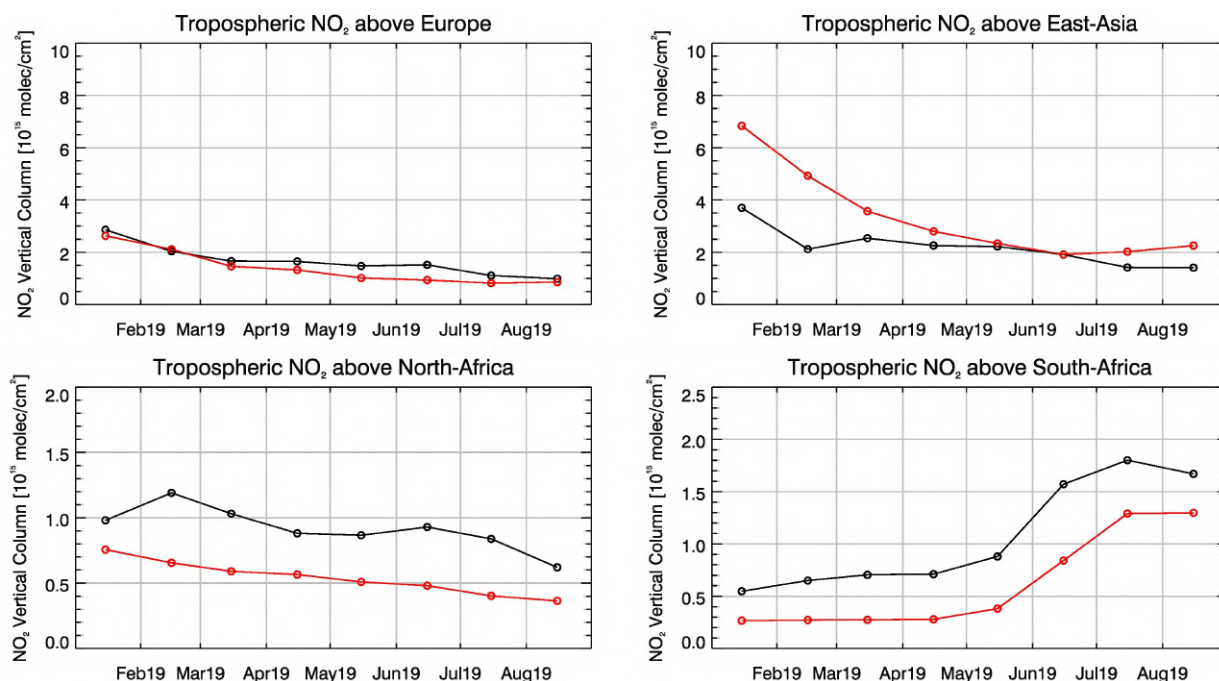
Figure 5.1.4: Time series of average tropospheric NO_2 columns [10^{15} molec cm^{-2}] from (black) TROPOMI compared to (red) o-suite model results for Jan - Aug 2019 (see Annex 2 for definition of regions). The upper panels represent regions dominated by anthropogenic emissions, and the lower panels represent those dominated by biomass burning.

Africa, the o-suite shows improved results since an update in July 2012 and the change to IFS-CB05 in September 2014. However, tropospheric NO_2 columns around December are still overestimated by the models. Summertime NO_2 columns over North-Africa are underestimated compared to the satellite data from 2015 onwards. The models (especially the o-suite) generally overestimate the seasonal cycle for South-Africa, particularly for 2014-2016 with an overestimation of the seasonal maximum which usually occurs around August (e.g. by a factor of 1.4 larger compared to GOME-2 retrievals in 2016). However, August maxima are in better agreement since the upgrade of the o-suite in 2017, but minima during SH summer remain underestimated.

Time series comparions between the o-suite and TROPOMI are shown in Figure 5.1.4 for Jan-Aug 2019. They show some differences with respect to the GOME-2 based ones: the o-suite overestimates values over East-Asia for Jan 2019 and underestimates values over the African regions during the whole period of Jan-Aug 2019 (the seasonal cycle over South-Africa is not overestimated in this case) according to the TROPOMI based comparisons. Differences in comparison results are in principal due to differences in observation time or differences in the retrieval products.

More NO_2 evaluation plots can be found on the CAMS website, see table 1.2.

5.2 Evaluation against ground-based DOAS observations

The NO_2 columns of the CAMS products have been compared with UVVIS DOAS profile measurements at Uccle and column data from the other stations.[1] This ground-based, remote-sensing instrument is sensitive to the NO_2 abundance in the lower troposphere, up to 1km altitude with an estimated uncertainty of 8%. Tropospheric NO_2 profiles and columns are validated (up to 3.5km or 10km). A description of the instruments and applied methodologies is the same all DOAS OFFAXIS measurements, see http://nors.aeronomie.be. It is important to mention here that the model partial column values are calculated from the smoothed model profiles. This guarantees that the model levels where the measurement is not sensitive do not contribute to the observed bias. We should mention that the measurement data is still catalogued as rapid delivery and not in the consolidated NDACC database.

Fig. 5.2.1 shows the biases for the validation periods Dec-Feb 2019 and June-August 2019 at the different sites. The biases have all increased, except at De Bilt where the bias changes sign (the bias was positive due to some extremely low values in the De Bilt measurement time series). At the urban sites at Uccle and Athens a strong underestimation is observed. For the other sites (Bremen, De Bilt and Cabauw) the o-suite analysis is able to capture only few of the high pollution events.

---

[1] No contribution from Xianghe, Reunion and OHP due to instrument failure.

Histogram plot of relative differences (M-O)/O for daily mean UVVIS NO2 timeseries 18DJF

Histogram plot of relative differences (M-O)/O for daily mean UVVIS NO2 timeseries 19JJA
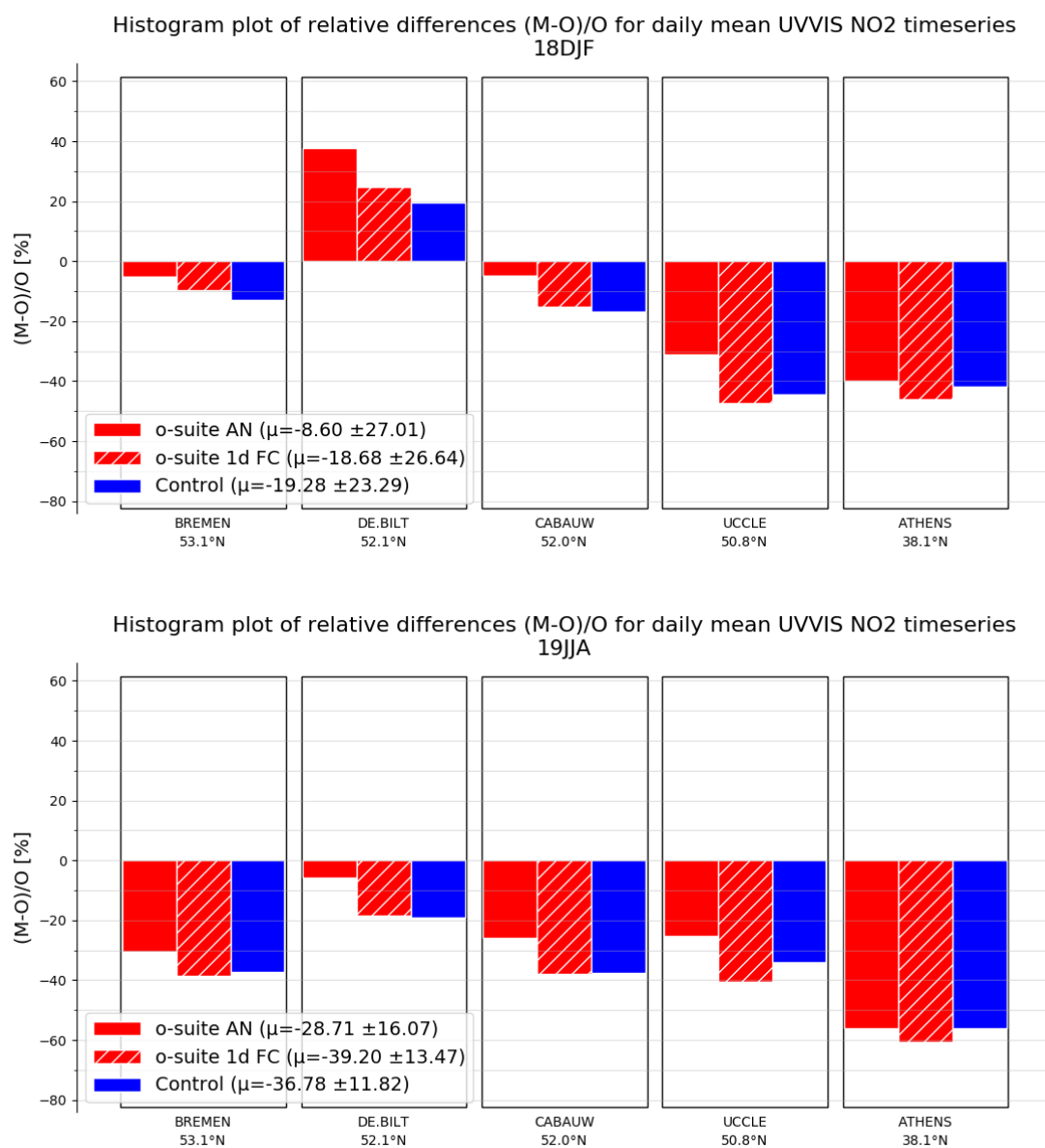
Figure 5.2.1: Table diagram showing the seasonal bias Dec-Feb 2019 (top) and June-August 2019 (bottom) for five stations, sorted by latitude. Compared to the validation period DJF, the relative biases in JJA have increased. AN means analysis, 1d FC means 1-day forecast.
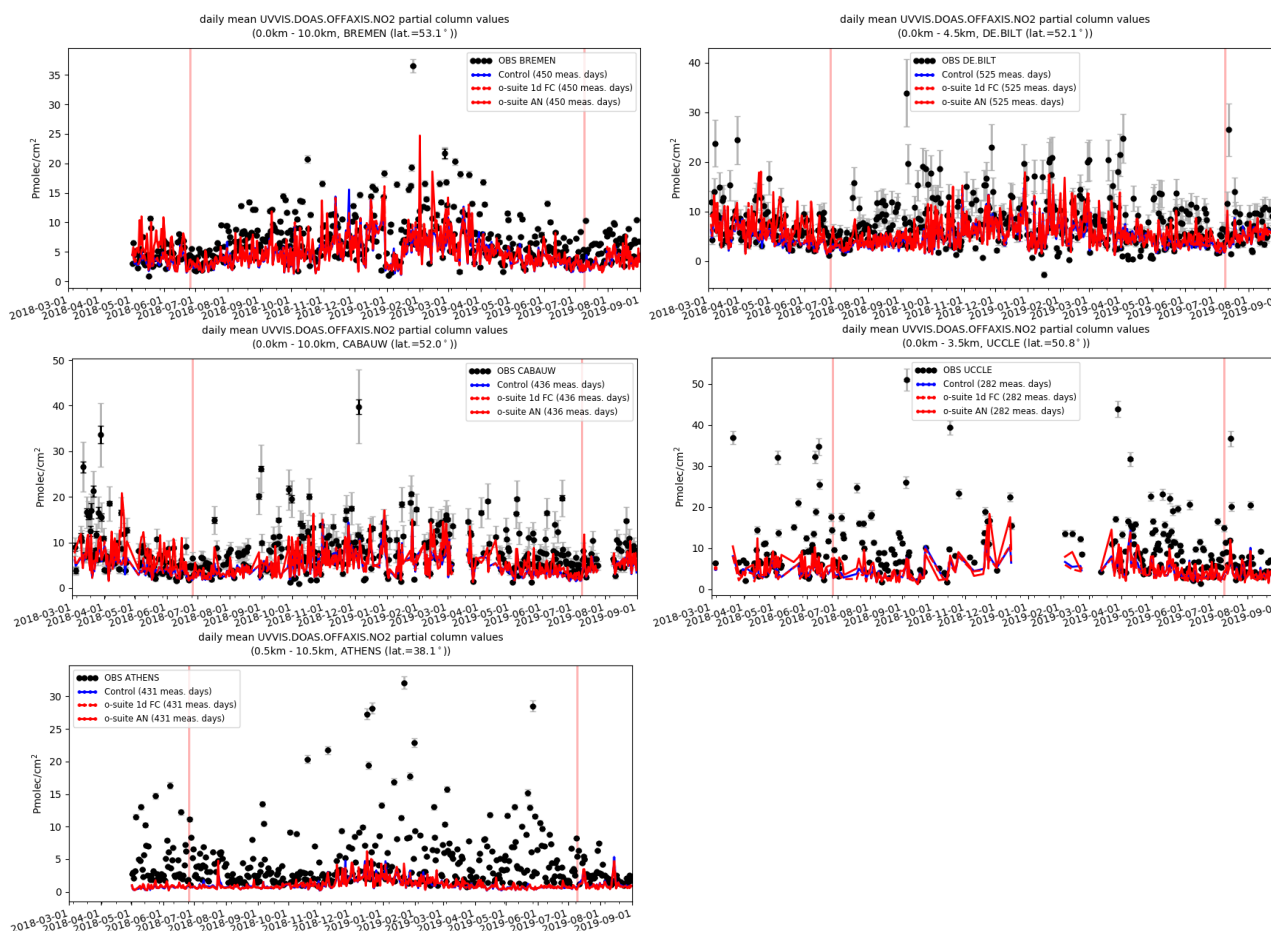
Figure 5.2.2: Time series of NO$_2$ partial columns at the five different sites. For all sites except Athens, background concentrations are well captured by the CAMS products. The o-suite and control product show little difference.

6. Formaldehyde

6.1 Validation against satellite data

In this section, simulations of tropospheric formaldehyde are compared to SCIAMACHY/Envisat HCHO satellite retrievals (IUP-UB v1.0) [Wittrock et al., 2006] for model data before April 2012 and to GOME-2/MetOp-A HCHO data (IUP-UB v1.0) [Vrekoussis et al., 2010] afterwards. As the retrieval is performed in the UV part of the spectrum where less light is available and the HCHO absorption signal is smaller than that of NO_2, the uncertainty of monthly mean HCHO columns is relatively large (20% – 40%) and both noise and systematic offsets have an influence on the results. However, absolute values and seasonality are retrieved more accurately over HCHO hotspots.

In Figure 6.1.1, monthly mean satellite HCHO columns are compared to model results for June 2019. The magnitude of oceanic and continental background values and the overall spatial distribution are well represented by the o-suite and control. The models overestimate values over regions in Central Africa which could be due to fire or biogenic emissions. As for tropospheric NO_2 values over boreal forest fires are overestimated over Siberia. However, the overestimation over boreal forest fires over Alaska and Canada does not show up in the HCHO comparisons, showing that the ability of the simulations to reproduce these values varies depending on the specie, location and time of the year.

Time series in Fig. 6.1.2 highlight three cases:

- East-Asia and the Eastern US, where HCHO is dominated by biogenic emissions. Model results and measurements generally agree rather well. However, all model runs underestimate the yearly cycle over East-Asia since 2012. In contrast to MOZART runs, MACC_CIFS_TM5 overestimated satellite values for the Eastern US since the middle of 2013. However, the newer IFS-CB05 runs perform well for Eastern US since 2015. For recent years and both regions, there is virtually no difference between the most recent o-suite run with IFS-CB05 chemistry and the corresponding control run without data assimilation. The variability or "ups and downs" in HCHO columns observed by GOME-2 since December 2014 is due to the lack of data (caused by instrument degradation) for these regions during winter in the Northern Hemisphere, leading to e.g. the negative values in the GOME-2 time series for Eastern US since December 2015. Summertime maxima are still underestimated over East-Asia despite of the higher resolution of the model runs since 2016.

- North-Africa, where biomass burning as well as biogenic sources largely contribute to HCHO and its precursors. Satellite observations over North-Africa tend to be slightly overestimated by IFS-CB05 chemistry model runs since 2014 and also the latest higher resolution model versions since July 2016. However, GOME-2 values are higher, and model values a bit lower this summer compared to previous years, resulting in a pronounced underestimation with respect to the satellite observations this summer.
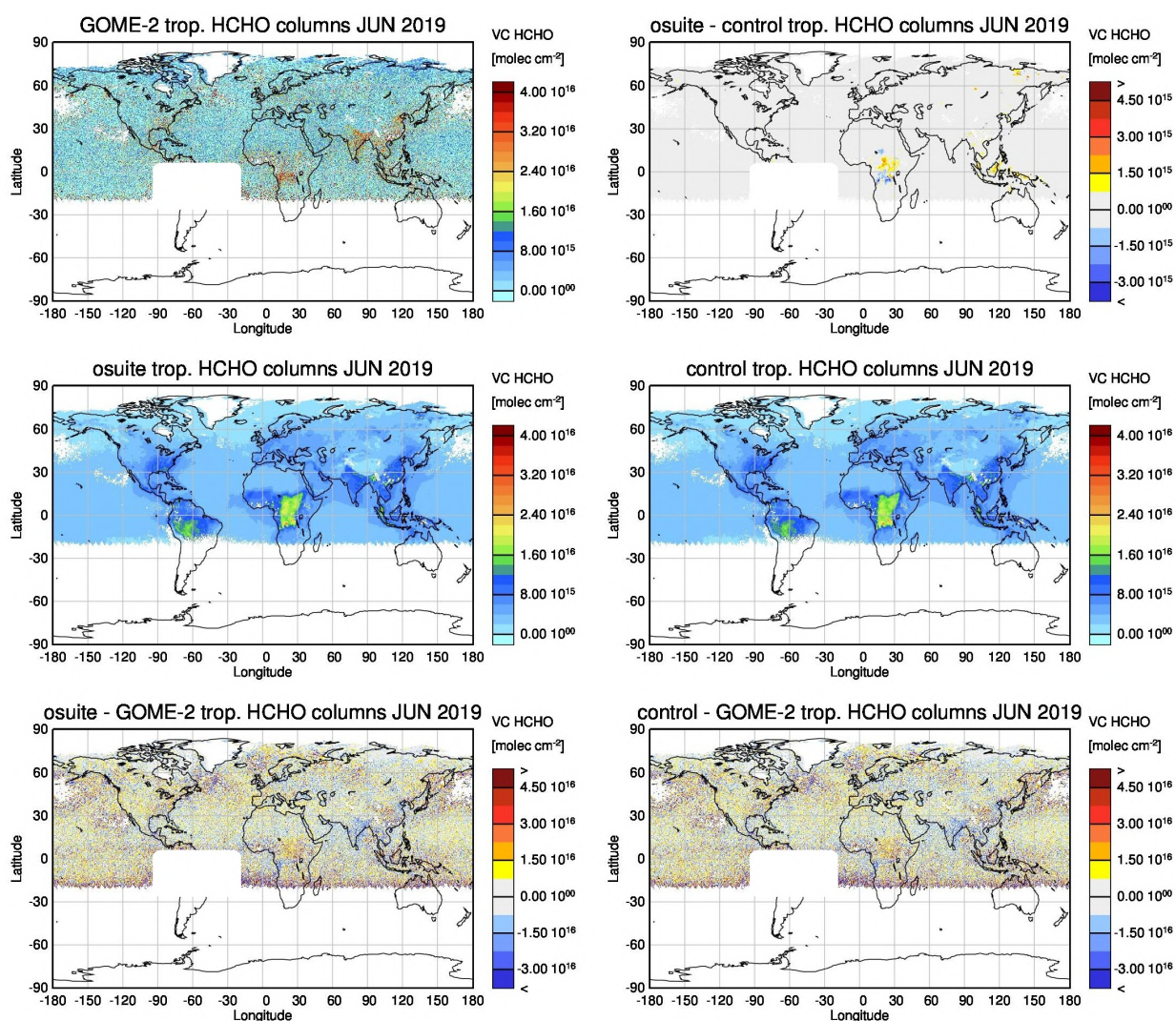
Figure 6.1.1: Global map comparisons of satellite-retrieved and model-simulated tropospheric HCHO columns [molec cm^{-2}] for June 2019. The top row shows monthly mean tropospheric HCHO columns retrieved by GOME-2, the second row shows the same but for model simulated averages. The third row shows differences of monthly means between models and GOME-2. GOME-2 data were gridded to model resolution (i.e. 0.4° x 0.4° degree). Model data were treated with the same reference sector subtraction approach as the satellite data. Satellite retrieved values in the region of the South Atlantic anomaly are not valid and therefore masked out (white boxes in all images except those which show model results only).

- Indonesia, where HCHO is also dominated by biogenic sources and biomass burning. Old MOZART based model versions generally overestimated satellite values here (by a factor of 3 – 4 in the second half of 2010) and failed to reproduce the observed seasonality. This may be due to the use of fire emissions including El Nino years, which experience much larger fire activities. MOZART simulations and observations agreed much better since late 2012. IFS-CB05 runs agree very well with satellite retrieved ones for December 2014 to August 2015. For September and October 2015, satellite retrieved HCHO columns show a pronounced maximum. 2015 was a strong El Nino year, which caused droughts and higher fire activity in
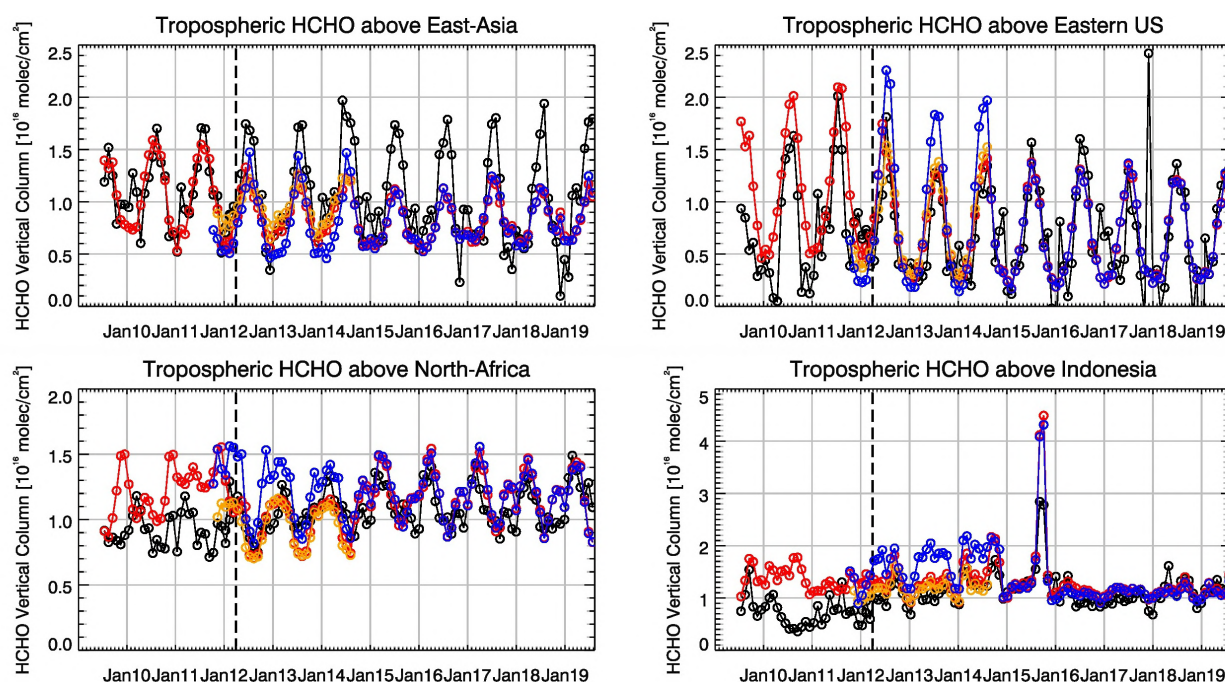
Figure 6.1.2: Time series of average tropospheric HCHO columns [1016 molec cm^{-2}] from SCIAMACHY (up to March 2012, black) and GOME-2 (from April 2012 onwards, black) compared to model results (red – o-suite, blue - MACC_fncrt_TM5/MACC_CIFS_TM5/control, orange - MACC_fcnrt_MOZ) for different regions. The blue line shows MACC_fcnrt_TM5 from November 2011 to November 2012, MACC_CIFS_TM5 results from December 2012 to August 2014 and control results from September 2014 onwards (the model run without data assimilation is termed control since Sep 2014). The regions differ from those used for NO2 to better focus on HCHO hotspots: East-Asia (25-40°N, 110-125°E), Eastern US (30-40°N, 75-90°W), Northern Africa (0-15°N, 15°W-25°E) and Indonesia (5°S-5°N, 100-120°E). Negative satellite retrieved values over Eastern US are due to a lack of data (caused by instrument degradation) during Northern Hemisphere winter months for this region. Vertical dashed black lines mark the change from SCIAMACHY to GOME-2 based comparisons in April 2012.

Indonesia. As for previous El Nino years, fire emissions used by IFS-CB05 seem to be largely overestimated, resulting in model-simulated HCHO columns which are almost twice as large as those retrieved by GOME-2. Further investigations (see previous reports) show that this is not caused by cloud flagging applied to the satellite and model data. There is mainly little variation from one month to another in both, satellite observations and model simulations since middle of 2016 and the magnitude of model and satellite values agrees overall well, the decrease in retrieved HCHO columns for Dec 17/ Jan 18 and an increase in May 2018 are not reproduced by the simulations.

For details on the HCHO evaluation: http://www.doas-bremen.de/macc/macc_veri_iup_home.html.

6.2 Evaluation against ground-based DOAS observations
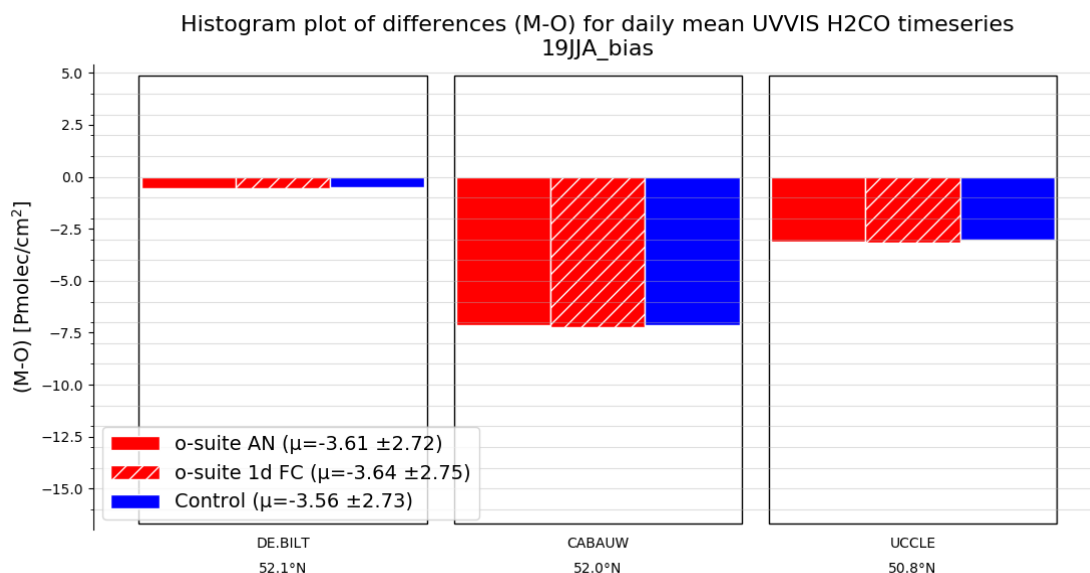
Figure 6.2.1: Table diagram showing the seasonal absolute bias in JJA 2019 for three stations, sorted by latitude.

In this section, we compare the H2CO columns of the CAMS products with UVVIS DOAS measurements at Uccle, Cabauw and De Bilt.[2] These ground-based, remote-sensing instruments are sensitive to the HCHO abundance in the lower troposphere. Tropospheric HCHO profiles and columns are validated (up to 3.5km (Uccle) or 10km (Cabauw and De Bilt)). A description of the instruments and applied methodologies is the same as for the MWR O_3 and FTIR O_3 and CO validations see http://nors.aeronomie.be. It is important to mention here that the model partial column values are calculated for the smoothed model profiles. This guarantees that the model levels where the measurement is not sensitive do not contribute to the observed bias. We should mention that the measurement data is catalogued as rapid delivery and not in the consolidated NDACC database.

Fig. 6.2.1 shows the absolute biases March - August 2019 at the different sites and indicates nearly vanishing bias for the different sites. At Cabauw, some high pollution events are not captured by the model and leads to a higher overall underestimation (Fig 6.2.2). From Fig. 6.2.1 and 6.2.2 we see little difference between the o-suite and the control run. Although the background column values are well captured by the products, the high emission events are not.

---

[2] No contribution from Reunion, Xianghe and OHP due to instrument failure.
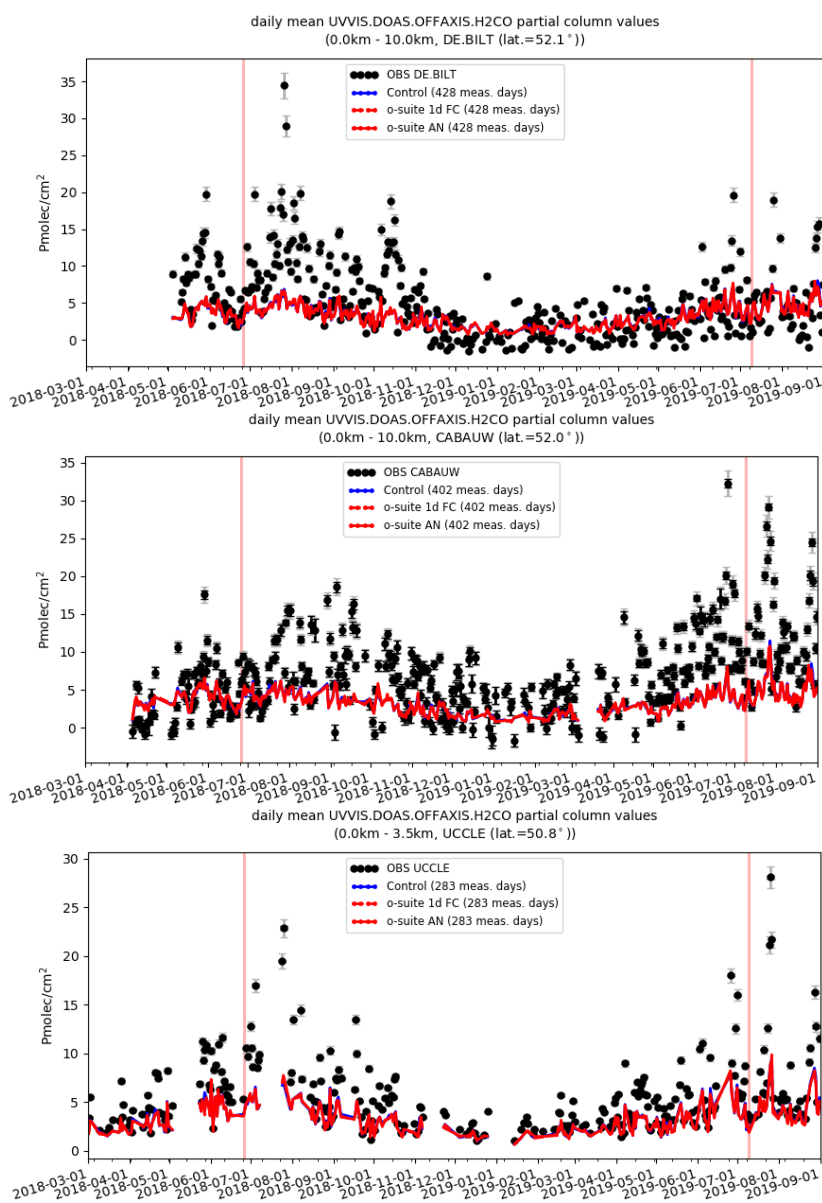
Figure 6.2.2: Time series of H2CO partial columns at the five different sites. All CAMS products underestimate the H2CO concentrations, except at De Bilt, where the model overestimates during the winter months.

7. Aerosol

7.1 Global comparisons with Aeronet and EMEP

The comparison of the CAMS simulation of time series of aerosol optical depth can be found for all Aeronet stations at: http://aerocom.met.no/cams-aerocom-evaluation/

More detailed evaluation including scores, maps, scatterplots, bias maps and histograms illustrating the performance of the aerosol simulation in the IFS system are made available through the AeroCom web interface. The model run can be compared here to e.g. the CAMS interim reanalysis and other models, such as the AeroCom Median model.

Correlation, based on daily aerosol optical depth and NRT Aeronet observations, has been rather stable recently. The o-suite forecast at +3 days shows only slightly lower correlation. See figure S3.

Part of the month-to-month variation in correlation is due to the varying quality and coverage of the Aeronet network. This has been improved by the version 3 from Aeronet. We use therefore version 3 level 1.5 for all global comparison to Aeronet.

The performance of the o-suite model exhibits some seasonal variation in AOD depending on region (Fig. 7.1.1). Since 45r1 there is an overall reduction of bias and the bias is more consistent across continents compared to previous versions. Noteworthy is the persistent AOD overestimation over North America (Fig. 7.1.1-bottom), but also a long-term trend to overestimation in East Asia. The latitudinal display of model and Aeronet AOD in the period investigated here (Fig. 7.1.2) shows a specific positive bias against Aeronet in the Southern Hemisphere.
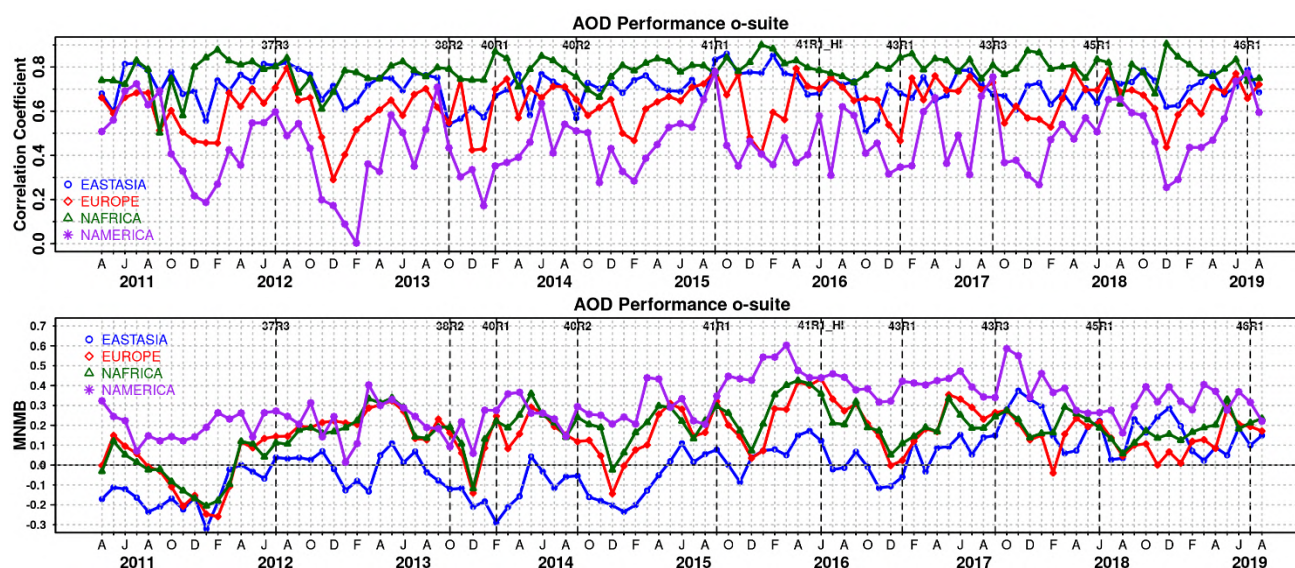
Figure 7.1.1. (top) Correlation coefficient and (bottom) modified normalized mean bias (MNMB) in AOD, since 2011, based on daily AOD comparison (Aeronet V3 level 1.5 data) in four world regions [East-Asia (blue); Europe (red); North Africa (green); North America (purple)] for the o-suite.
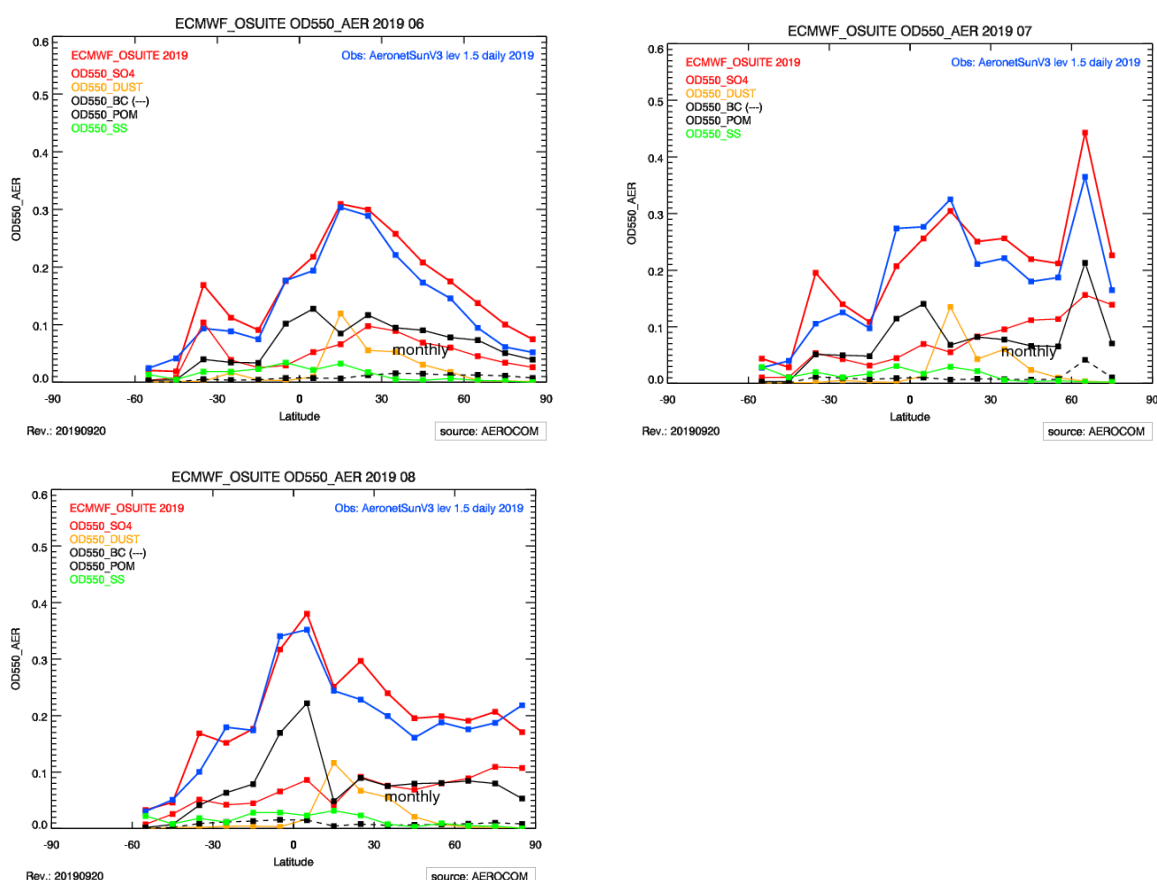
Figure 7.1.2. Aerosol optical depth of o-suite (red) compared to latitudinally aggregated Aeronet V3 level 1.5 data (blue) for the three months covered by this report.
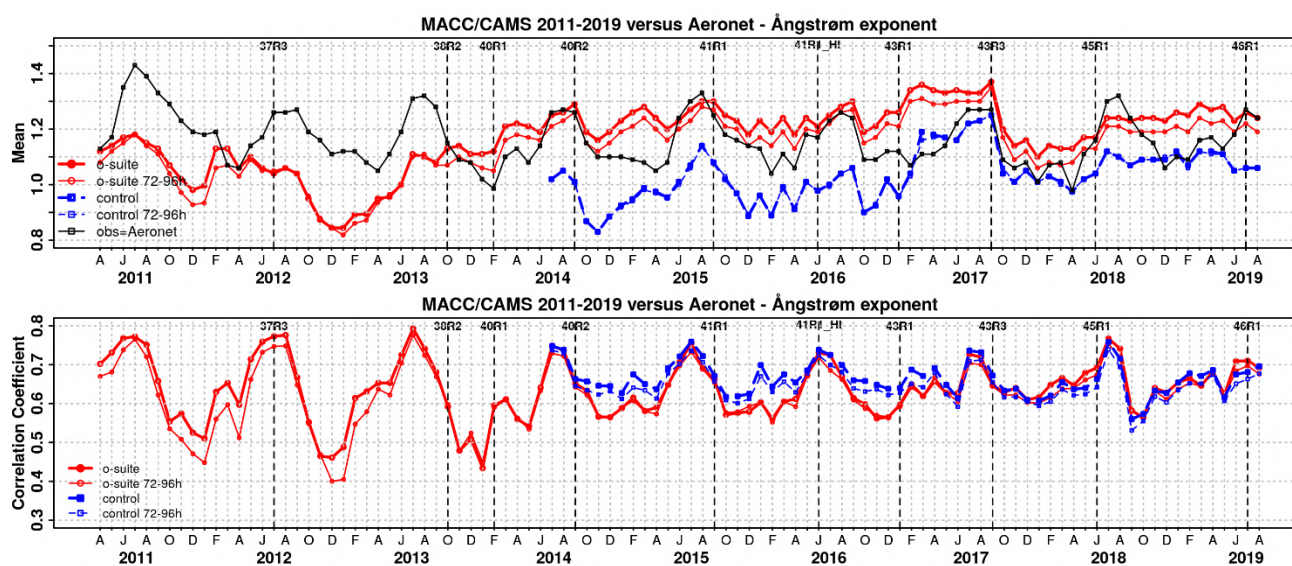
Figure 7.1.3. a) (top) Evolution of mean Ångström exponent in o-suite and control at Aeronet sites (Aeronet V3 level 1.5 data), based on matching monthly mean values. o-suite (thick red curve); o-suite at last forecast day (light red curve); control (blue dashed curve); control at last forecast day (light blue dashed curve). b) (bottom) Correlation using daily matching Ångström exponent.
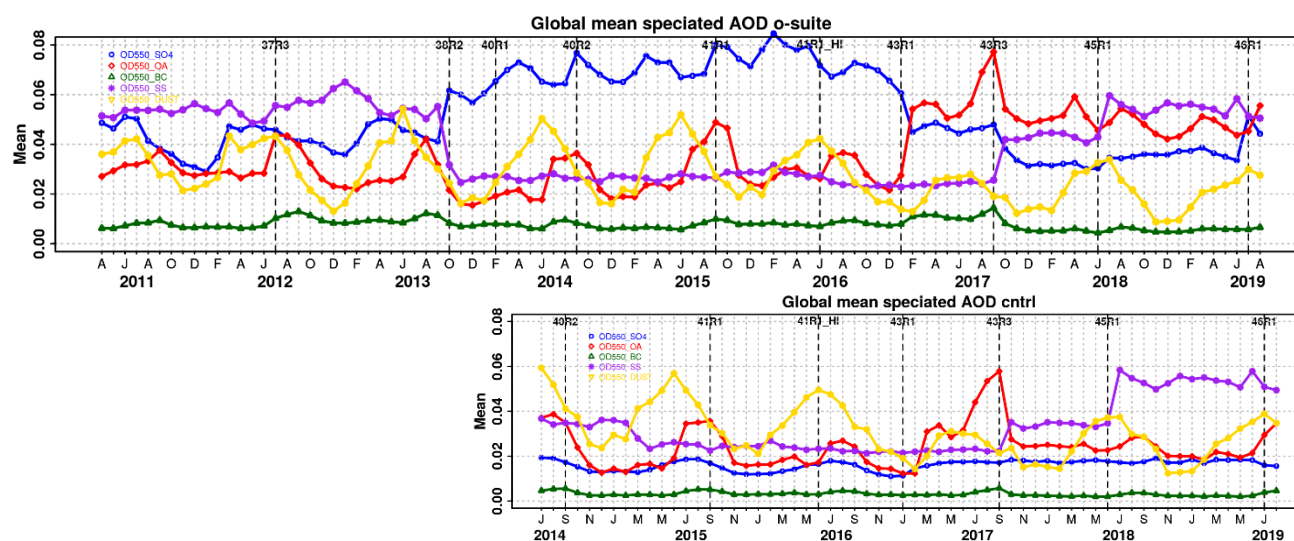
Figure 7.1.4. Evolution of the aerosol components of total AOD@550nm [OD550_SO4 = sulphate(blue); OD550_OA = organics(red); OD550_BC = black carbon(green); OD550_SS = sea salt(purple); OD550_DUST = dust(yellow)] in o-suite and control simulation.

o-suite				
	Mean MAM 2019 0-24h	Change wrt to first day on day 4	Mean JJA 2019 0-24h	Change wrt to first day on day 4
AOD@550	0.160	-19%	0.181	-11%
BC-OD@550	0.005	-29%	0.006	-16%
Dust-OD@550	0.026	2%	0.028	10%
OA-OD@550	0.054	-29%	0.048	-23%
SO4-OD@550	0.032	-24%	0.043	-22%
SS-OD@550	0.043	-14%	0.053	-3%

Table 7.1.1. Mean global total and speciated AOD in the o-suite for the last two periods covered by the VAL report and change after 3 forecast days.

The simulated aerosol size distribution may be validated to first order using the wavelength dependent variation in AOD, computed as Ångström exponent, with higher Ångström exponents indicative of smaller particles. We find in JJA 2019 a small bias (Figure 7.1.3-a). Temporal and spatial variability is difficult to capture, but correlation from all daily data is lower than for AOD (Figure 7.1.3-b and S3). Figure 7.1.4 shows that the Sep 2017 and Jun 2018 model changes are responsible for a shift in Ångström exponent. More organic matter seems to shift the size distribution to smaller sizes. The model upgrade in February 2017 with a bugfix for sea salt and improved parameterisations for SO_4 lead to sea salt increased with 45% while sulphate further decreased a bit. Sea salt has increased further due to a new sea salt emission scheme implemented in the June 2018 model upgrade and is back to earlier 2011-2013 levels. The Grythe et al. (2014) parametrization introduced in CY45R1 is supposed to increase emissions over tropics due to dependency on SST, and this confirmed by comparing a test run with 45R1 in 2017 with the 2017 o-suite, which is low in sea salt. Since the latest model upgrade in July 2019 with the improvement to the sulphur cycle, the SO_4 seem to have increased to same levels as before the Feb 2017 upgrade.
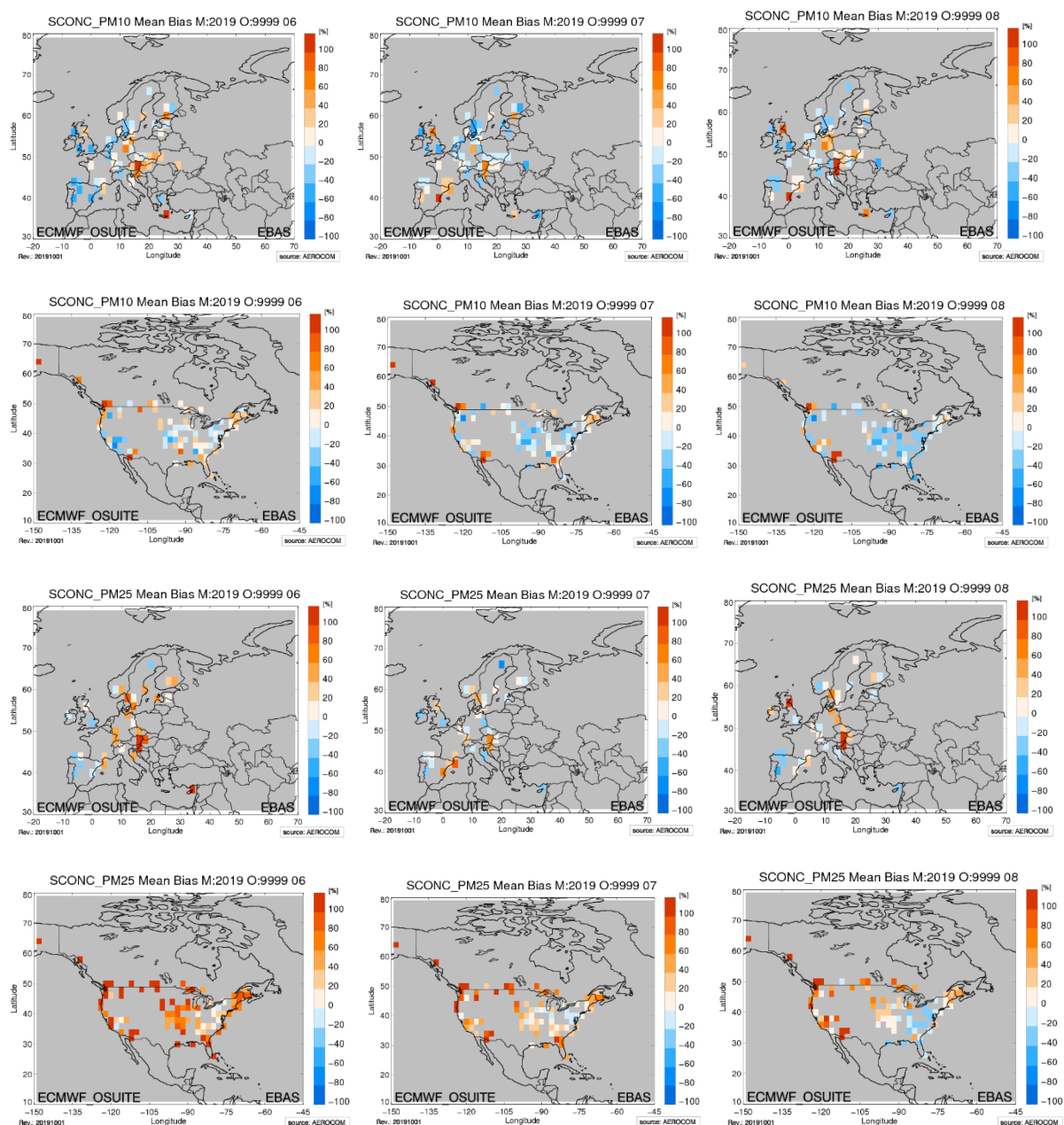
Figure 7.1.5. Bias [%] map of monthly mean PM10 and PM2.5 concentrations at EMEP (Europe, first and third row) and IMPROVE sites (North America, second and fourth row) for June (left column), July (middle) and August 2019 (right); simulated o-suite versus EMEP/IMPROVE derived climatological average (2000-2009).

The o-suite uses data assimilation to obtain an analysis of the aerosol field. In the forecast period, however, a-priori model parameterisations and emissions (except fire emissions, which are kept in the forecast equal to the latest GFAS emission values) determine increasingly the aerosol fields. The performance of the day three forecasted AOD fields as compared to the first guess is shown in Figure S3 in the summary of this report. Table 7.1.1 shows an average global decrease in total aerosol optical depth during the first four forecast days, dominated by sulphate and organics. The

Figure 7.1.6. Temporal evolution of monthly mean average PM25 and PM 10 concentrations at EMEP (Europe) and IMPROVE sites (North America) and data fraction within a factor 2 of observed; ca 160 sites, observed data averaged from data available in EBAS from 2000-2009.
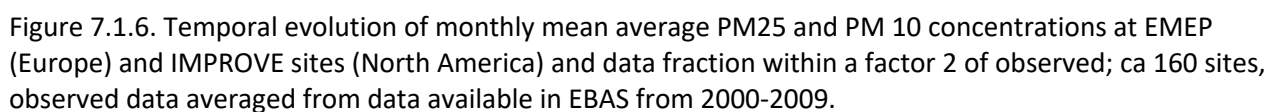
control run with no assimilation shows somewhat less AOD (-11% compared to o-suite, see figure S3). All this supports the conclusion that either a-priori IFS aerosol and aerosol precursor sources are too small, or sinks are too effective in the IFS model.

Surface concentration of particulate matter below 10 µm (PM10) and below 2.5 µm (PM25) from the o-suite experiment have been validated against data from 160 background IMPROVE and EMEP stations. A climatological average has been constructed from data in the period 2000-2009 as available in the EBAS database hold at NILU. The data availability is not the same at all stations, and sometimes covers only a few years.

A negative MNMB bias of PM10 in Europe and an overestimate in North America PM2.5 appears (Fig. 7.1.5), consistent with the AOD bias in the two regions. Figure 7.1.6 shows the evolution of mean observed and simulated PM10 and PM2.5. The biggest change appeared in July 2017 with the bias of o-suite now becoming positive overall for PM25 and also for PM10, except for the period autumn-spring 2018/19. Shown is also the statistics of being within factor 2, a more robust metrics for a comparison to climatological data. This statistical indicator has clearly improved over time, indicating best PM10 and PM25 performance in summer months for the o-suite. O-suite is also better most of the times than the control simulation for PM10. For PM25 the difference is less clear, but since September 2017 (upgrade to 43R3) the control is performing better than the o-suite, except since latest upgrade in July 2019 the o-suite seem to perform as good as the control.

7.2 Dust forecast model inter-comparison: Validation of DOD against AERONET, and comparisons with Multi-model Median from SDS-WAS

The 72-hour forecasts (on a 3-hourly basis) dust aerosol optical depth (DOD) from CAMS o-suite and control have been validated for the period 1 June 2019 – 31 August 2019 against AERONET Spectral Deconvolution Algorithm (SDA) cloud-screened observations, MODIS/Terra and Aqua Collection 6.1 Level 3 (1º x 1º) and SDS-WAS Multi-model Median DOD. The SDS-WAS Multi-model Median DOD is obtained from (currently) twelve dust prediction models participating in the Sand and Dust Storm Warning Advisory and Assessment System (SDS-WAS) Regional Center for Northern Africa, Middle East and Europe (http://sds-was.aemet.es/). At those sites where the SDA products are available, the dust AOD evaluation will be complemented with AOD-coarse which is fundamentally associated to maritime/oceanic aerosols and desert dust. Since sea-salt is related to low AOD (< 0.03; Dubovik et al., 2002) and mainly affects coastal stations, high AOD-coarse values are mostly related to mineral dust.

During this season, MODIS (Figure 7.2.1) shows that major dust activity in Northern Africa (seasonal AOD up to 0.3) is concentrated in latitudes south of 20º N with high seasonal enhancements (seasonal AOD over 0.7). In North Africa, both CAMS products simulate the main areas of dust activity in comparison with MODIS (see Figure 7.2.1). Both CAMS experiments underestimate the dust activity over the Bodélé (in Chad) and overestimate dust in the source region between Mali-Algeria-Niger borders as well as Sudan, Iraq and North Saudi Arabia. The CAMS o-suite shows lower season values (seasonal DOD up to 0.9) than control (seasonal DOD up to 1.2), which are in general higher than the SDS-WAS multi-median product (seasonal DOD up to 0.5). The CAMS o-suite reproduces the dust transport over the North Atlantic region although the maximum dust activity is
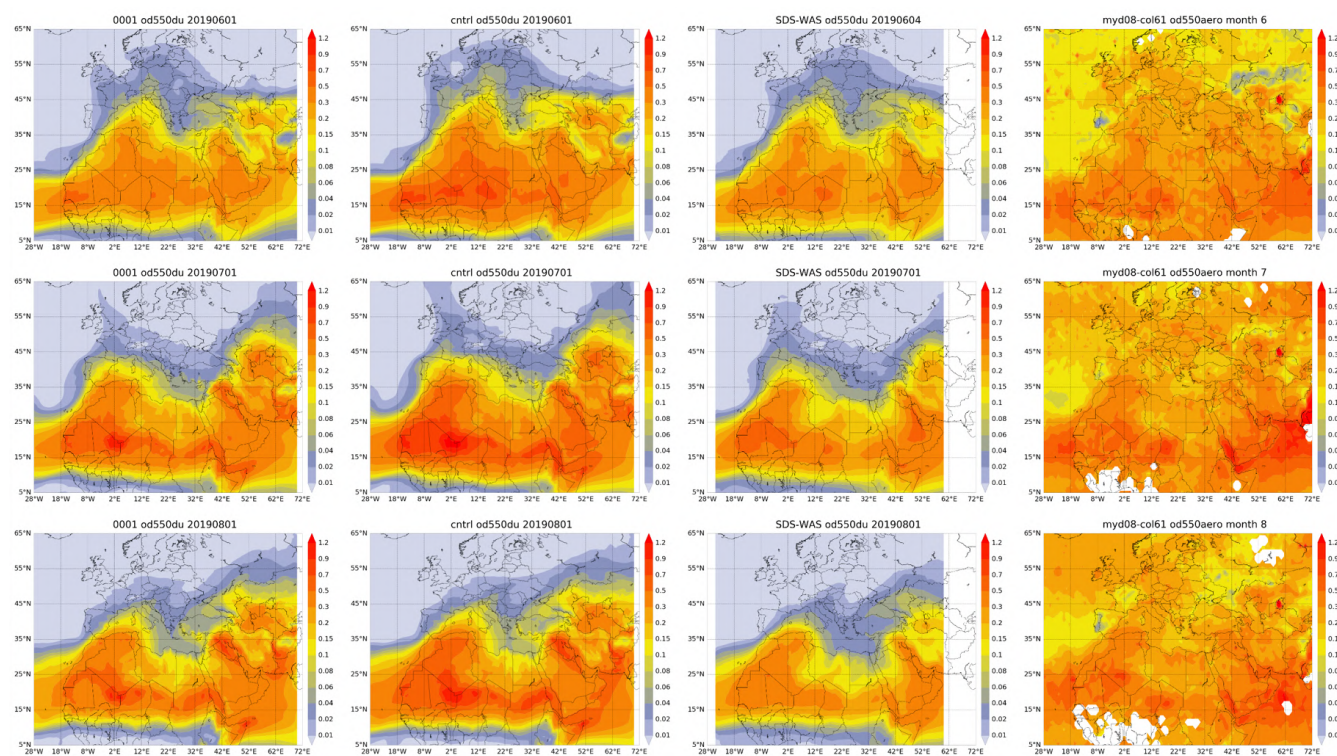
Figure 7.2.1: Averaged DOD 24h forecast from o-suite (first column) and control (second column), DOD of the multi-model SDS-WAS Median product (third column) as well as AOD from MODIS/Aqua Collection 6.1 Level 3 combined Dark target and Deep Blue product (last column) for the study period. Top row: June. Middle row: July. Bottom row: August 2019.

shifted to Mali, Niger and Algeria border and Eastern Sahara instead of Chad as shown in MODIS. Also, DOD over Iraq and in the Mediterranean Basin appears overestimated in the comparison with MODIS. These changes in dust activity in the main source regions are linked to the new dust module implemented in the operational CAMS model since early-July 2019.

From June to August, the o-suite reproduces the daily variability of AERONET direct-sun observations (see Figure 7.2.2a and Table 7.2.1) with a correlation coefficient of 0.82, averaged over all the AERONET sites (as in the case of the SDS-WAS multi-model product). This is similar to the control experiment with a correlation coefficient of 0.83. Regarding the mean bias (MB), both CAMS products (o-suite and control) slightly overestimate the AERONET observations, resulting in a MB of 0.00 for o-suite and 0.05 for control, in contrast to the SDS-WAS multi-model that underestimates (MB of -0.03). Similar results are obtained in the comparison with the AERONET SDA observations (see Figure 7.2.2b).

The impact of the upgrade in early July can be observed in Fig. 7.2.1. For the o-suite, the DOD results over the NAMEE, show that the correlation drops from 0.86 in June to 0.81 in August and the mean bias increases from -0.04 in June to 0.03 in August. This is an overestimation compared with the results of the SDS-WAS ensemble, which shows a correlation of 0.85 and a mean bias of -0.01.
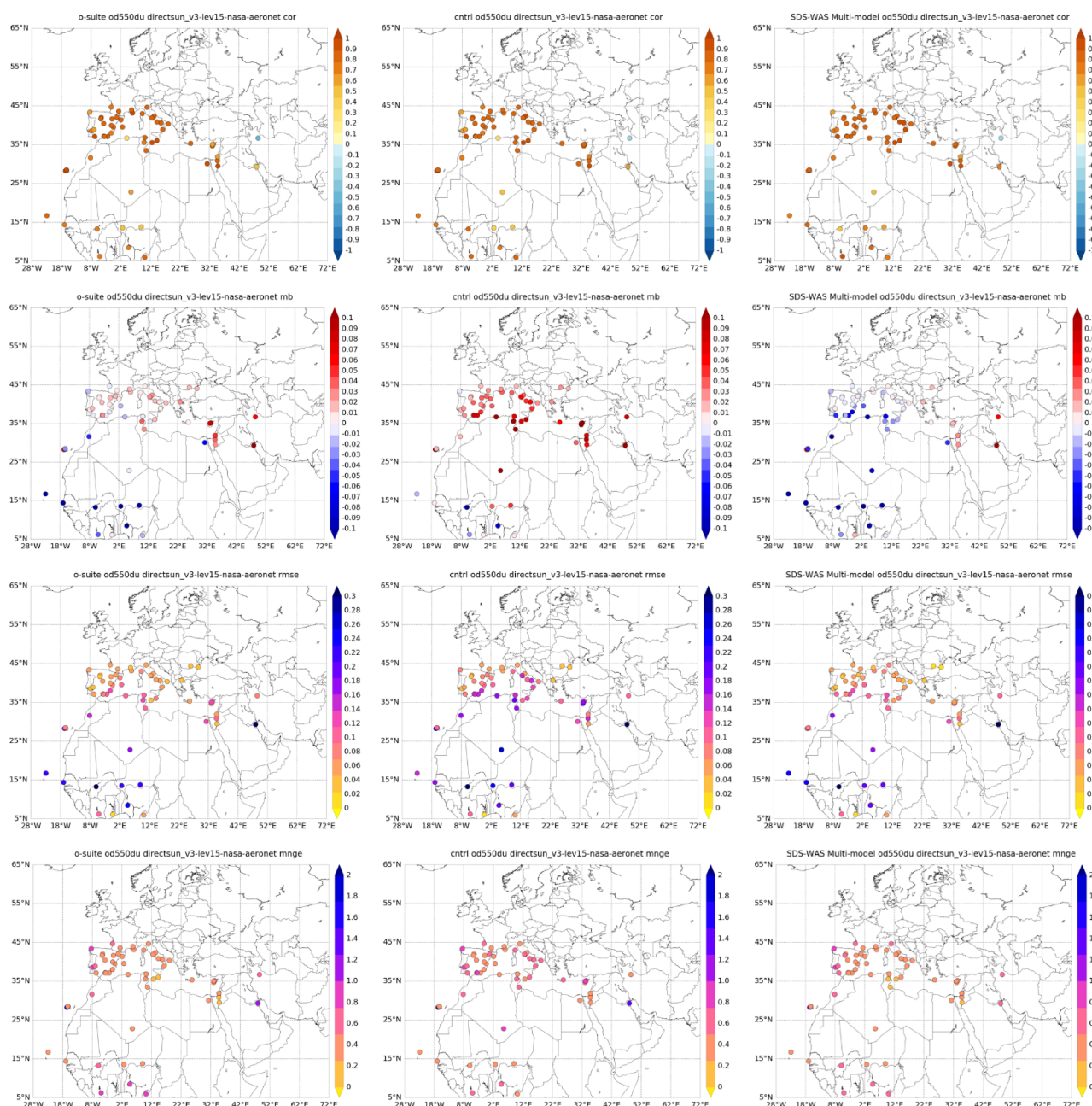
Figure 7.2.2a: Skill scores (correlation coefficient, MB, RMSE and FGE) for 24-hour forecasts of CAMS o-suite (left column), control (central column) and DOD Multi-model SDS-WAS Median (right column) for the study period. Dust-filtered AOD from AERONET direct-sun is the reference.

Over desert dust sources in the Sahara and the Middle East (see Table 7.2.1 as well as Tamanrasset INM and Kuwait University AERONET sites in Figure 7.2.3a), o-suite and control can reproduce the daily variability with correlation coefficients of 0.57 and 0.64 for o-suite and 0.51 and 0.66 for control, respectively for Sahara and the Middle East. Overestimations are observed in both CAMS experiments over the Sahara (MB of 0.17 for control and 0 for o-suite) and the Middle East (MB of 0.39 for control and 0.33 for o-suite). These results are closer to the SDS-WAS Multi-model results
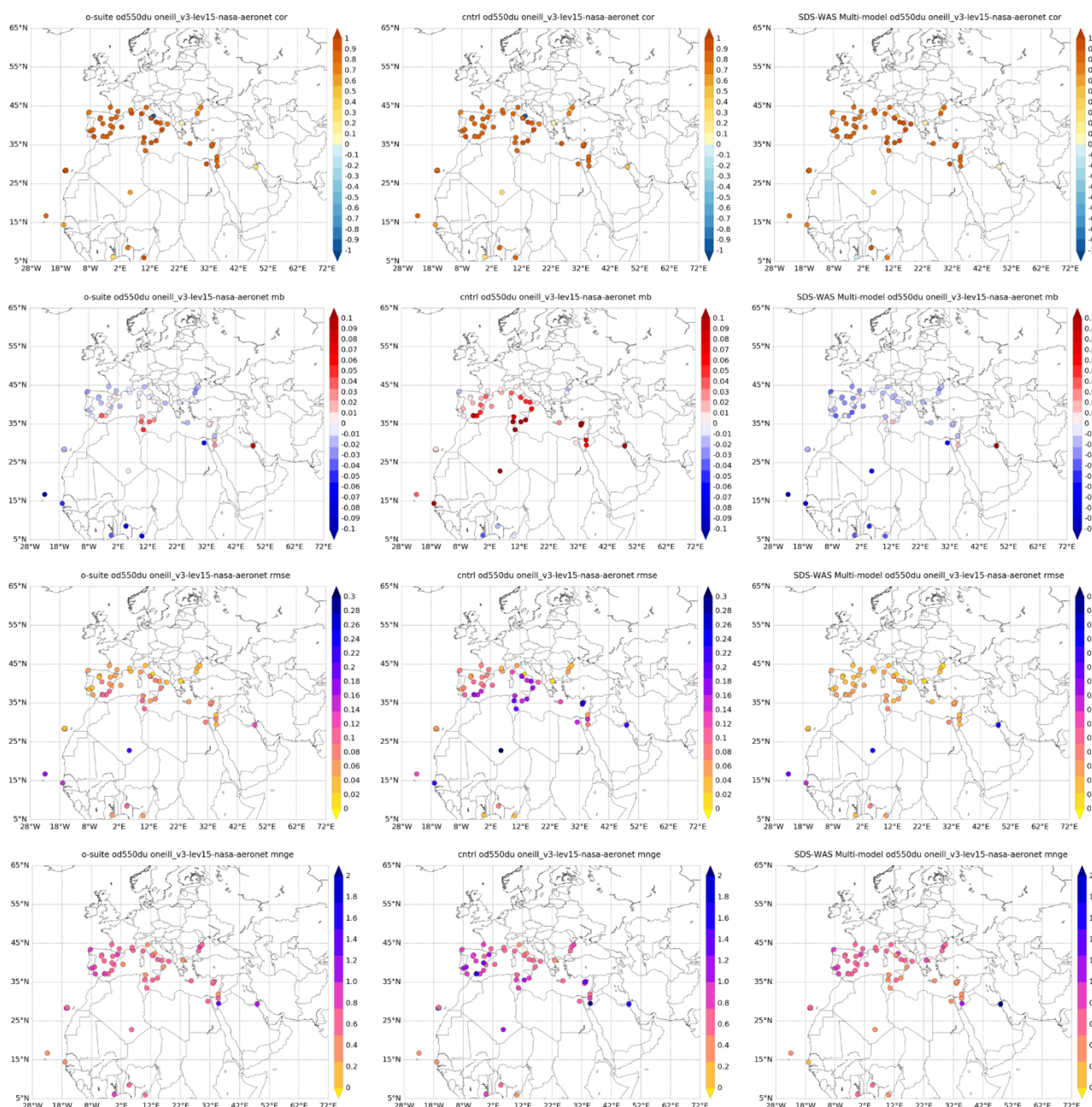
Figure 7.2.2b: Skill scores (correlation coefficient, MB, RMSE and FGE) for 24-hour forecasts of CAMS o-suite (left column), control (central column) and DOD Multi-model SDS-WAS Median (right column) for the study period. AOD-coarse from AERONET SDA is the reference.

for Sahara (seasonal correlation of 0.55 and MB of -0.08) and better for the Middle East (seasonal correlation of 0.71 and MB of 0.20). As shown in Figure 7.2.3a, the higher CAMS overestimations (that are reduced in the o-suite experiment) are observed from mid-July after which there is a pronounced change in the behaviour of the CAMS model (see Kuwait University in Figure 7.2.3a). As indicated before, these changes in the dust activity in the main source regions are linked to the new dust module implemented in the operational CAMS model that became operational on the 9th of July 2019.

Table 7.2.1: Skill scores (MB, FGE, RMSE and r) of 24h forecasts for CAMS o-suite, CAMS control and SDS-WAS Multi-model Median for the study period, and the number of data (NDATA) used. Dust AOD (DOD) from AERONET is the reference.

	NDATA	control				o-suite DOD				SDS-WAS Median DOD			
		MB	FGE	RMSE	r	MB	FGE	RMSE	r	MB	FGE	RMSE	r
Sahara	297	0.17	0.42	0.26	0.51	0.00	0.02	0.16	0.57	-0.08	-0.23	0.17	0.55
Sahel	822	0.01	0.14	0.21	0.60	-0.14	-0.28	0.23	0.64	-0.14	-0.22	0.22	0.73
Tropical North Atlantic	139	-0.01	0.00	0.14	0.76	-0.17	-0.41	0.21	0.81	-0.19	-0.45	0.24	0.77
Subtropical North Atlantic	569	0.02	0.48	0.10	0.78	-0.01	0.30	0.07	0.82	-0.02	0.17	0.08	0.80
North Western Maghreb	332	0.02	-0.31	0.17	0.74	-0.05	-0.54	0.15	0.74	-0.13	-0.79	0.17	0.77
Western Iberian Peninsula	678	0.01	1.74	0.06	0.60	0.01	1.73	0.04	0.61	0.00	1.71	0.04	0.61
Iberian Peninsula	1216	0.03	1.73	0.06	0.89	0.01	1.68	0.05	0.86	-0.01	1.64	0.06	0.87
Western Mediterranean	2697	0.04	1.36	0.10	0.84	0.01	1.25	0.08	0.84	-0.02	1.19	0.09	0.87
Central Mediterranean	2767	0.07	1.29	0.14	0.88	0.02	1.16	0.08	0.89	-0.02	1.07	0.08	0.90
Eastern Mediterranean	906	0.08	1.61	0.14	0.83	0.03	1.50	0.10	0.79	0.00	1.51	0.08	0.89
Eastern Sahara	-	-	-	-	-	-	-	-	-	-	-	-	-
Middle East	47	0.39	1.39	0.51	0.66	0.33	1.30	0.46	0.64	0.20	1.19	0.27	0.71
All sites	**10470**	**0.05**	**1.19**	**0.14**	**0.83**	**0.00**	**1.04**	**0.11**	**0.82**	**-0.03**	**0.98**	**0.11**	**0.87**

In the Sahel (see Figure 7.2.2 and Table 7.2.1), the o-suite shows strong underestimations (MB of -0.14) in comparison with control (with MB of 0.01). However, the o-suite better reproduces the observed daily variability (with a correlation value of 0.64 for o-suite in comparison to 0.60 for control). The underestimations observed in the o-suite in the Sahel are also spread to the Tropical North Atlantic (MB of -0.17 for the o-suite, see Dakar in Figure 7.2.3b and Table 7.2.1).

Over long-range transport regions (see Figure 7.2.2), the CAMS o-suite model results are better than control over the sub-Tropical North Atlantic region with correlation values from 0.78 for control to 0.82 for o-suite (see Table 7.2.1). Over the Iberian Peninsula and the Mediterranean, both CAMS products show high correlations (between 0.79 and 0.89) and overestimations (MB between 0.01 and 0.03 for o-suite and between 0.01 and 0.08 for control).

The Central Mediterranean region has the best results in the AERONET comparison in terms of correlation, see Table 7.2.1 and Tunis Carthage in Figure 7.2.3b. Both experiments can reproduce the daily variability with correlation coefficient of 0.89 and 0.88 for both o-suite and control, respectively. The Western Iberian Peninsula presents slightly lower correlation values (0.60 for control and o-suite). This is related the minimum dust activity during this period in this region. In the case of the North-Western Maghreb, both CAMS experiments can reproduce the daily variability with a correlation coefficient of 0.74. Meanwhile, the control run is overestimating the observations (MB of 0.02) and these overestimations are reduced in the o-suite (MB of -0.05). In the Eastern Mediterranean, o-suite correlation coefficient is 0.79, smaller than for control (which is 0.84). The higher overestimations observed in control are reduced in the o-suite (see Table 7.2.1 and Ben Salem in Figure 7.2.3b).
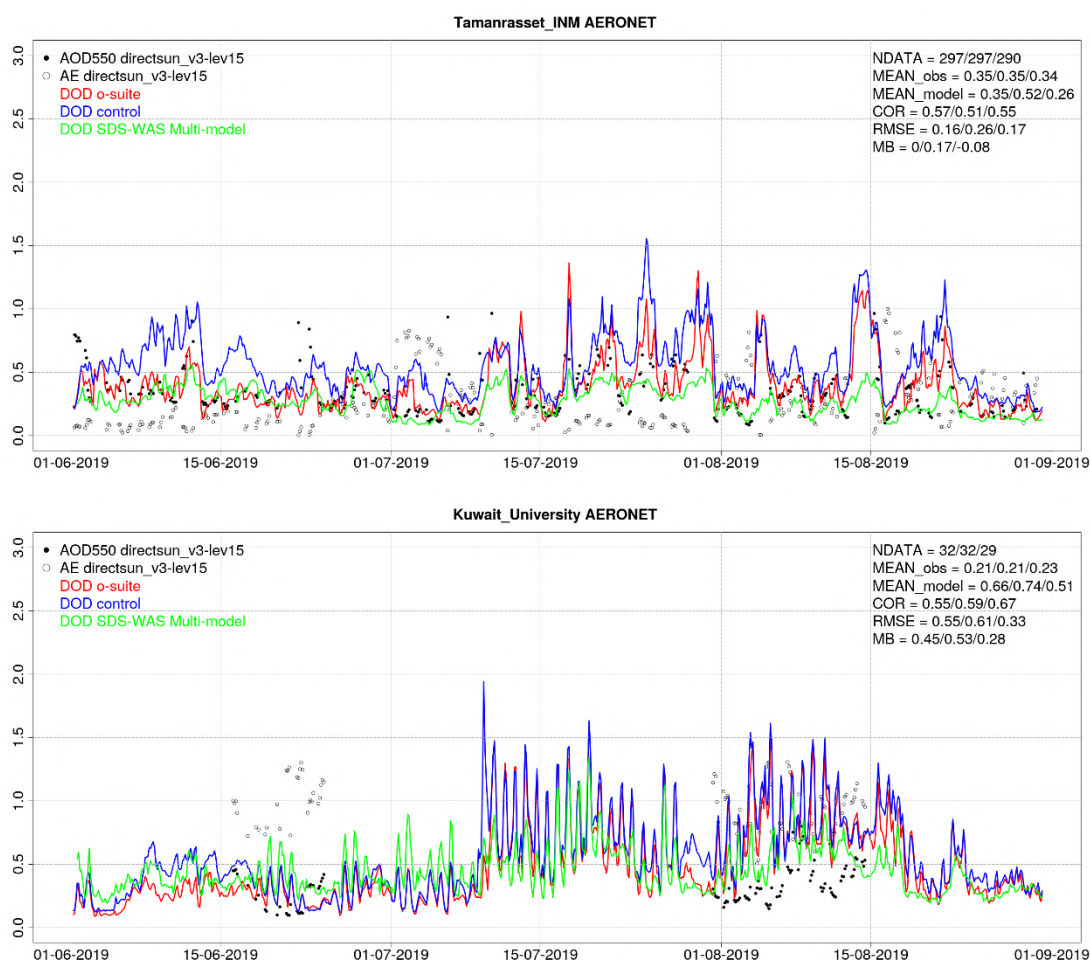
Figure 7.2.3a: AOD and Angstrom Exponent from AERONET Direct-sun (black dots), DOD o-suite (red line), DOD control (blue line) and DOD Multi-model SDS-WAS Median (green line) for the study period over Tamanrasset-INM (Sahara) and Kuwait University (Middle East). Skill scores per each individual site and model (o—suite/control/ SDS-WAS Multi-model) are shown in the upper right corner (NDATA: available 3-hourly values used for the calculations, MEAN observations, MEAN model, COR, RMSE, MB).

The comparison of the 1-day and 3-day forecasts shows that the prediction skill is stable during the forecasts in comparison with AERONET direct-sun observations, with correlation coefficients of 0.82 (0.83), 0.81 (0.83), and 0.79 (0.81) respectively for 24, 48 and 72h forecasts for all the sites for o-suite (control).
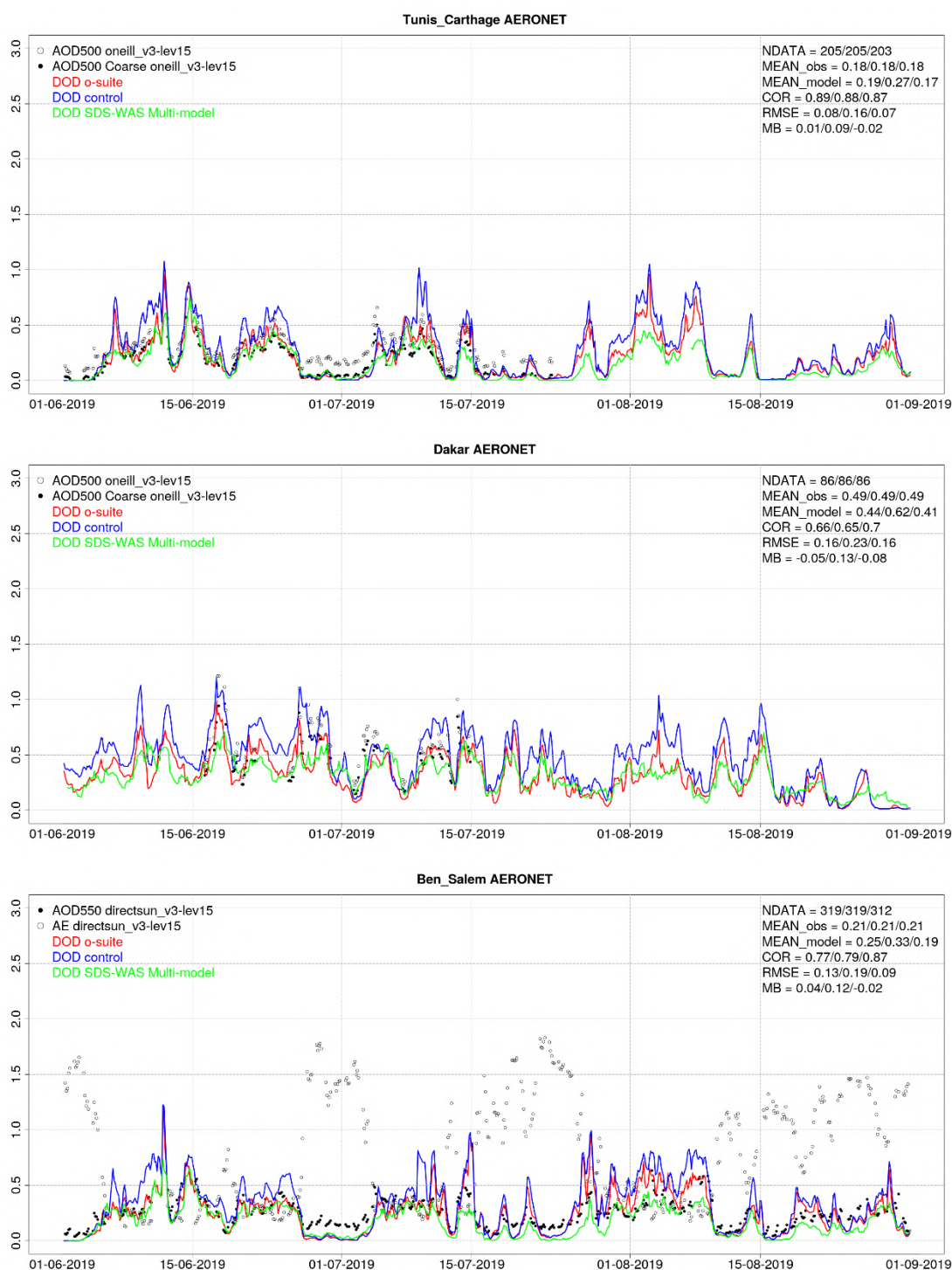
Figure 7.2.3b: AOD and AOD-coarse from AERONET SDA (black dots), DOD o-suite (red line), DOD control (blue line) and DOD Multi model SDS-WAS Median (green line) for the study period over Tunis-Carthage (North Western Magrebh), Dakar (Tropical North Atlantic) and Ben Salem (Eastern Mediterranean) Skill scores per each individual site and model (o—suite/control/SDS-WAS Multi-model) are shown in the upper right corner (NDATA: available 3-hourly values used for the calculations, MEAN observations, MEAN model, COR, RMSE, MB).

7.3 Backscatter profiles

The technical specifications of the ceilometer data, evaluated parameters and methods are described in the report CAMS-84 D8.1. In this section, the temporal and vertical variation of the backscatter coefficient (BSC) profiles are evaluated statistically as the bias, correlation, and standard deviation of o-suite '0001' and control run 'gzyh' and 'h7c4' (as of 10 July 2019) versus ceilometers and are summarized in Taylor plots.

A model upgrade on 10 July 2019 increased the number of vertical levels from 60 to 137 (Fig. 7.3.1). Nitrate NO_3 (two size bins) and ammonia NH_4 are included since the upgrade. Both are coupled to the gas-phase nitrate chemistry, and also for sulphur species the chemistry and aerosol schemes have been coupled for consistency. Anthropogenic and biogenic emission inventories, anthropogenic SOA production, biomass burning injection and the dust emission scheme were updated to produce larger particles.
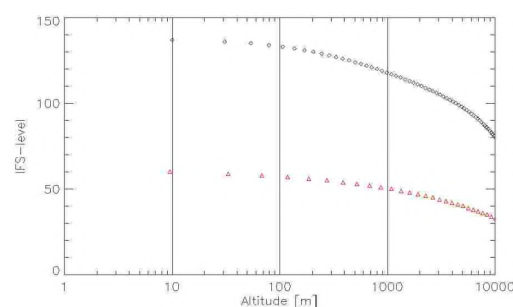
Figure 7.3.1: 60 vs 137 C-IFS levels

Period Overview

In Figure 7.3.2, showing the maximum AOD over Germany, the model aerosol optical depth (AOD) indicates a period with higher SO_4 concentrations starting with the update on 10 July, while all other components show no marked changes. Saharan dust events occurred from 9 - 16 June and from 24 June - 2 July. With respect to AOD, all aerosol components follow their usual seasonality.
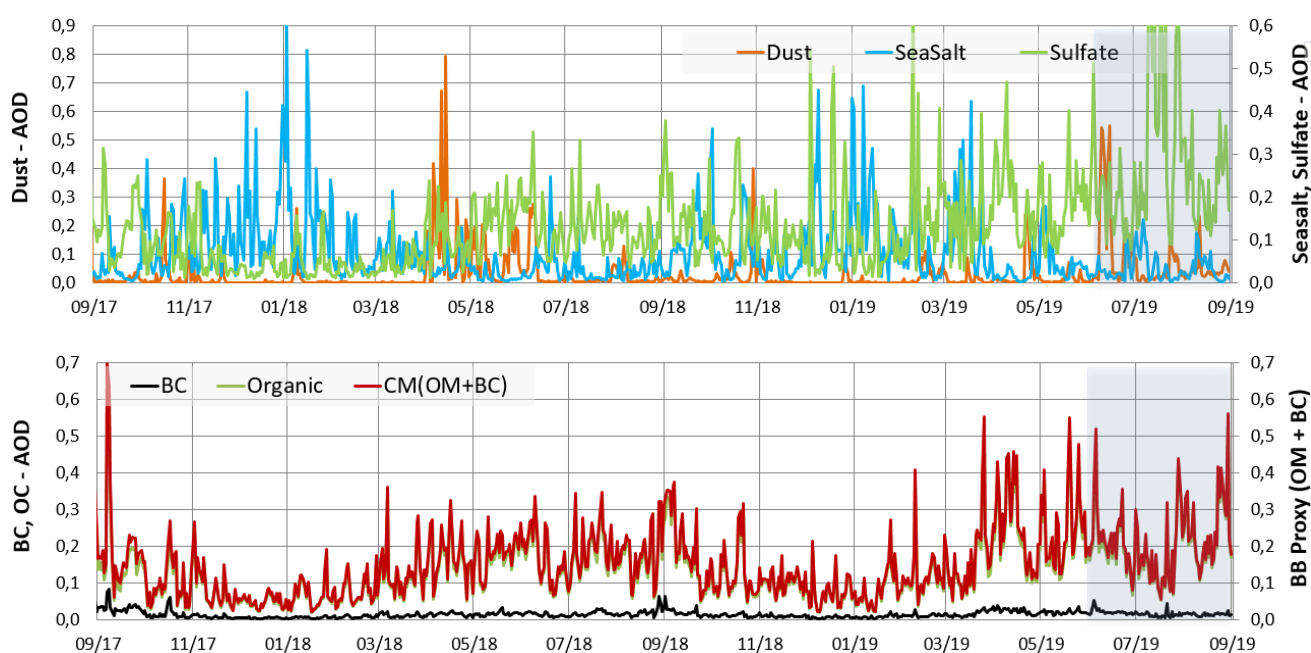
Figure 7.3.2: Maximum daily AOD over Germany for aerosols included in the IFS model from 09/2017 - 08/2019: sea salt (blue), dust (orange), sulphate (light green), black carbon (BC, black), organic matter (green), proxy for 'biomass burning' (as OC+BC - red). Note the different y-axes for the aerosol species.
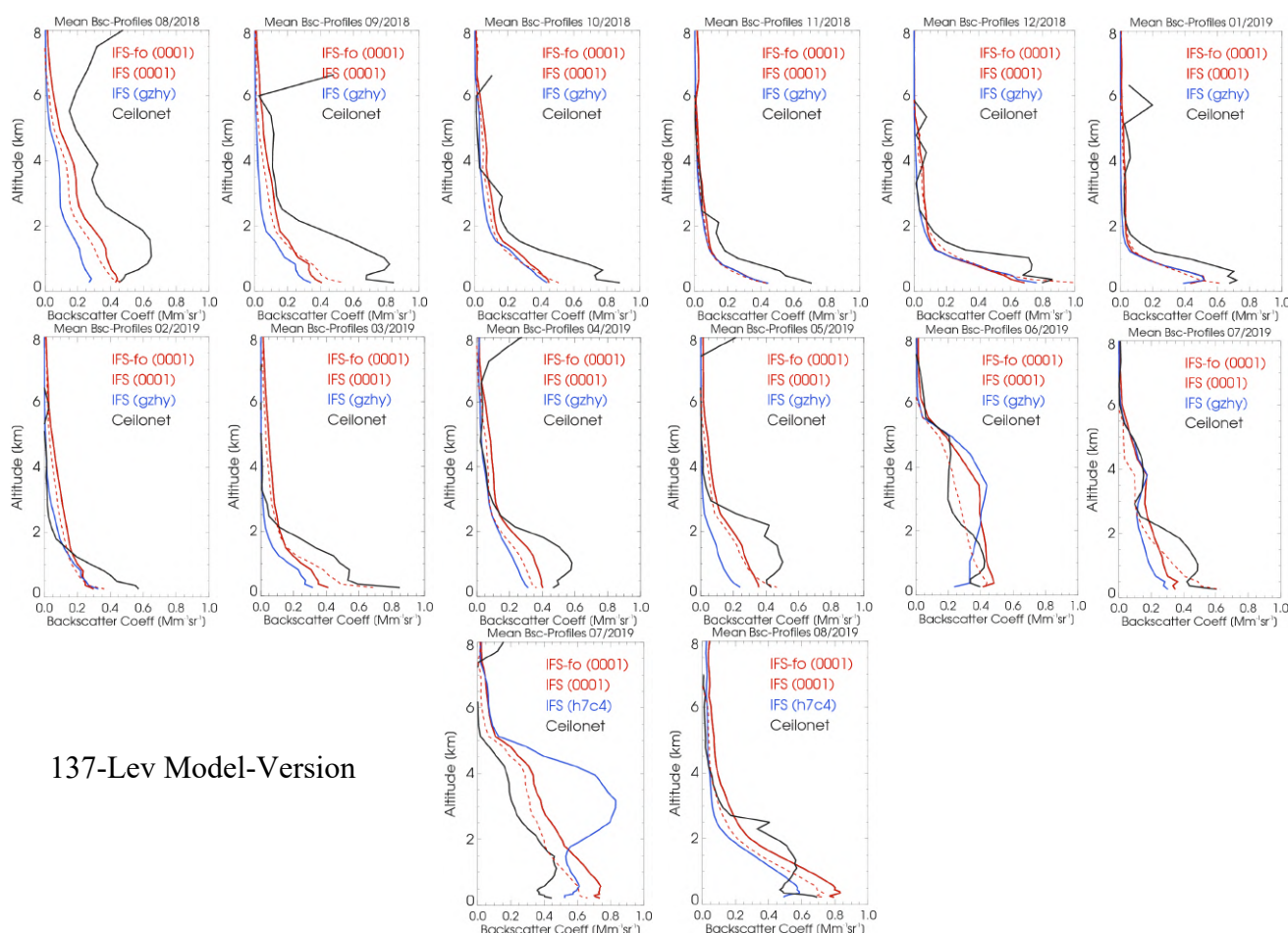
137-Lev Model-Version

Figure 7.3.3: Monthly mean profiles of backscatter coefficients from o-suite (red), control run gzhy/h7c4 (blue), and ceilometers (black) combined from 21 German stations in Aug 2018 to Aug 2019. The profiles are partly contaminated by remaining cloud artefacts.

Mean profiles:

In the new IFS cycle 46r1, implemented on 9 July 2019, nitrate NO_3 and ammonia NH_4 have been added and, likewise sulphate SO_4, have been coupled w.r.t homogeneous (gas-phase) and heterogeneous (particle phase) chemical processes. They contribute roughly 10-30% of aerosol mass in the rural central European PBL, as neutralized forms NO_3NH_4 or $(NH_4)_2SO_4$. Simultaneously, emissions of most aerosol components were significantly upgraded, which makes some diligence necessary to infer the individual changes' contributions to model performance.

Given only about 1.5 months of new data and that some important (though not grave) adaptions of our code are outstanding, only preliminary conclusions can be drawn (see remark below). The low bias of model BSC in the PBL has clearly decreased, but the profile is still strongly smoothed, i.e. the step at the top of the PBL is not captured notably better with 137 levels than with 60 levels (51L instead of 27L <8 km altitude), same for o-suite and control (Fig. 7.3.3). Effectively, the assimilation adds aerosol mass but also vertically redistributes it from the PBL to the FT. However, the amplitude of the model vertical profile, coded in the standard deviation in the Taylor plots has now approached closer to the observations (reference).
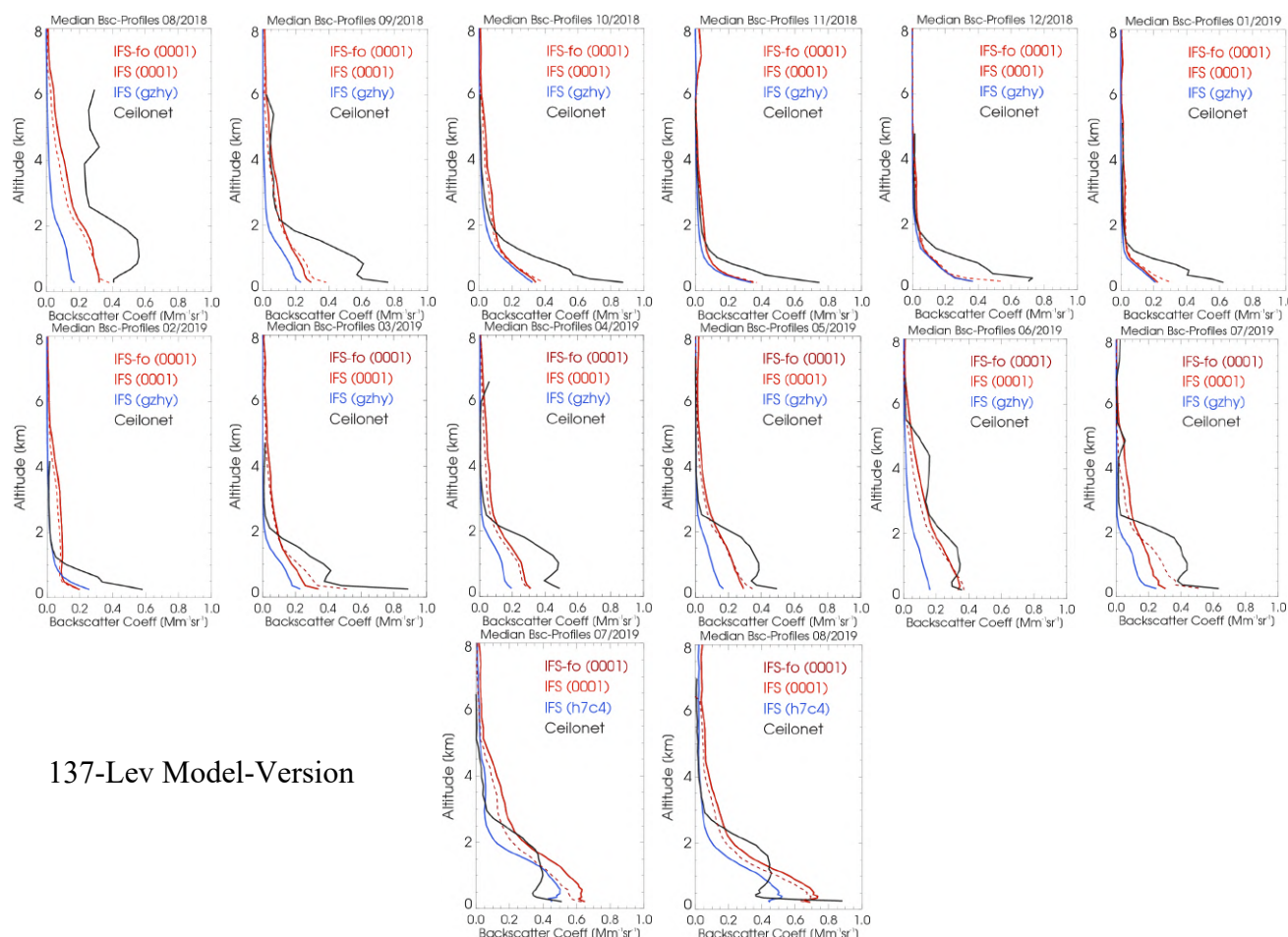
137-Lev Model-Version

Figure 7.3.4: Monthly median profiles of backscatter coefficients from o-suite (red), control run gzhy/h7c4 (blue), and ceilometers (black) combined from 21 German stations in Jun 2018 to Aug 2019. The profiles are partly contaminated by remaining cloud artefacts.

The June/July control run profiles reflect the influence of Saharan dust events SDE (17 days with $AOD_{SD}>0.05$) which have typically been over-estimated by the model. The vertical amplitude is thus larger in the mean profiles (more affected by events) than in the median profiles, which reflect more background conditions.

It must be noted that the profiles 'IFS (0001)' are calculated preliminarily using mass-to-backscatter conversion coefficients from SO_4 for both NO_3 and NH_4, which is a rough approximation for the moment and will be replaced by correct values in the next report. The dashed profiles using the IFS-forward operator are, however, correct and confirm that this approximation does not change the main conclusion concerning the progress seen by the model upgrade in July 2019.
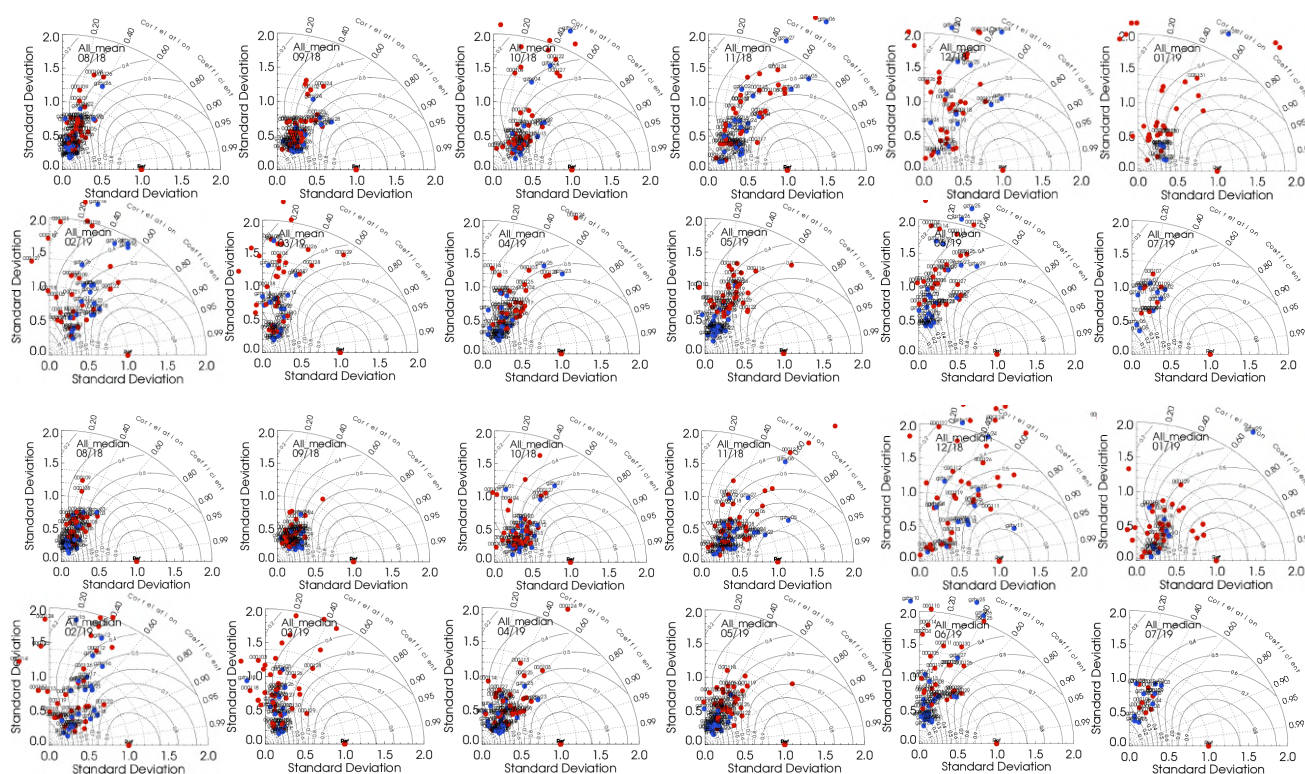
Fig. 7.3.5: Taylor polar plots with daily average standard deviation of vertical profiles vs correlation coefficient, averaged over 21 German ceilometer sites for Aug 2018 – Jul 2019 (O-suite red, control blue). Top: mean values, Bottom: median values.

Taylor Plots (60-level version only):

The average coefficient of correlation between modelled and observed vertical backscatter profiles clusters around r = 0.2-0.6 (Fig. 7.3.5). The absolute standard deviation (SD) are normalized to the SD of observations per day, as reference value at SD ≡ 1.0. In summer 2018, the o-suite (red dots) performs better than the control run (blue dots) in terms of SD (profile amplitude), which becomes less marked during autumn and winter but re-appears in summer 2019. There is a large day-to-day and also a seasonal variation of the performance. In June 2019 the spread is due to Saharan dust events which are overestimated by the model, causing higher amplitude of the profiles by strong layers and thus higher standard deviation. On the other hand small vertical displacements cause a low correlation coefficient, although the SD plumes are reproduced at geographically truthful positions.

Summary

The preliminary conclusion concerning changes due to the model upgrade in July 2019 are:

- The low bias of model BSC in the PBL has clearly decreased.

- The high bias of model BSC in the FT has not changed notably.

- The profile is still strongly smoothed, i.e. the step at the top of the PBL is not captured notably better with 137 levels than with 60 levels (51L instead of 27L <8 km altitude), same for o-suite and control.

- The amplitude of the model vertical profile, dominated by the PBL, has approached closer to the observations (reference).

- Daily averages of Pearsons correlation coefficients in June/July (run gzhy) cluster around 0.2-0.6 and exhibit large day-to-day variability due to Saharan dust events, where the profile amplitude is better but the correlation is worse due to vertical displacement of dust layers.

7.4 Aerosol validation over Europe and the Mediterranean

Three-hourly aerosol optical depth (AOD) and surface concentration (PM10 and PM2.5) from the o-suite experiment and control experiment have been validated for the period 1 June 2019 – 31 August 2019 against AERONET direct-sun cloud-screened observations.

Aerosol optical depth over the Mediterranean

The CAMS o-suite can reproduce the daily variability of AERONET direct-sun observations. In Western, Central and Eastern Mediterranean, the correlation coefficient decreases from 0.82, 0.86 and 0.80, to 0.74, 0.82 and 0.65, respectively for control and o-suite during summer (see Figure 7.4.1). Both CAMS experiment overestimated the AERONET observations in the Mediterranean Basin in control (MB of 0.11, 0.17 and 0.19 for Western, Central and Eastern Mediterranean regions respectively) and o-suite (MB of 0.13, 0.16 and 0.19 for Western, Central and Eastern Mediterranean regions respectively) as it is shown in Figure 7.4.1. The highest peaks on CAMS AOD simulations are linked to desert dust intrusions occurring during the whole season In the whole Mediterranean Basin as it is shown in Barcelona (Spain), IMAA Potenza (Italy) and Finokalia (Crete) AERONET sites (see Figure 7.4.2) achieving AOD values up to 1 in Barcelona.

Surface aerosol concentrations over Europe

For summer, PM10 and PM2.5 results of CAMS o-suite and control show similar skill scores in comparison with EIONET-Airbase observations (see Figure 7.4.3). CAMS model tends to overestimate the PM10 and PM2.5 EIONET-Airbase observations in Central Europe with MB up to 10µg/m^3 (see Figure 7.4.3 and Figure 7.4.4) while the PM10 and PM2.5 observed values are underestimated at Iberian Peninsula and North Atlantic sites. Overall, for all the EIONET-Airbase sites, o-suite presents higher overestimations in PM10 and PM2.5 (with MB of 2.20 and 3.45 µg/m^3, respectively) than control (with MB of -1.11 and 1.11 µg/m^3, respectively).

The upgrade of the CAMS model is clearly identified in the EIONET-Airbase time series (see Figure 7.4.5). From 9th July, in southern European sites, PM10/PM2.5 ratio is higher than in previous periods indicating an increase of the coarse particles at surface levels. These changes in the size distribution is associated to the upgrade in the dust module in the CAMS model, which has shifted dust size distributions to coarse fraction increasing the coarse fractions at surface levels.
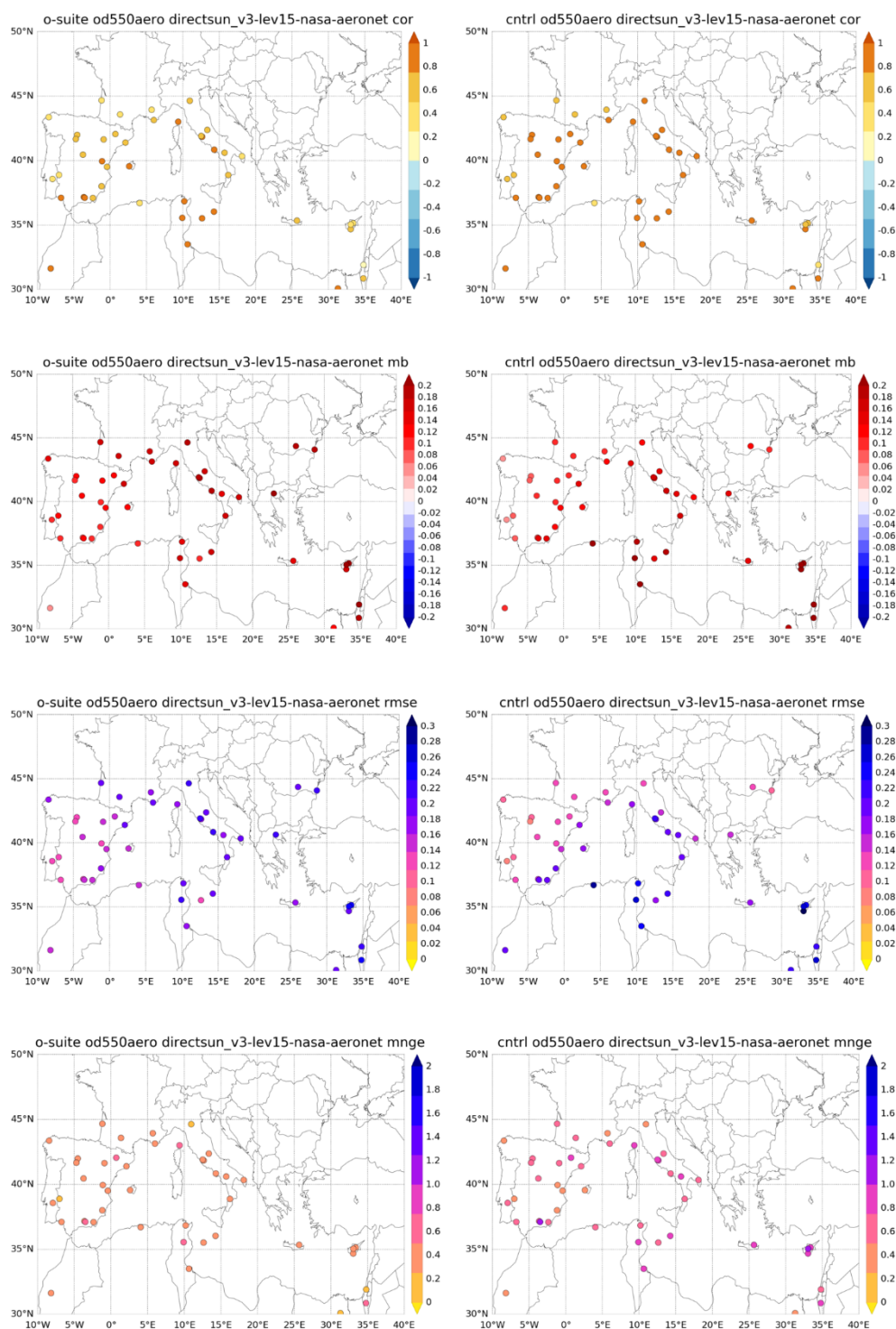
Figure 7.4.1: Skill scores (correlation coefficient, MB, RMSE and FGE) for 24-hour forecasts of CAMS o-suite and control for the study period. AOD from AERONET direct-sun is the reference.
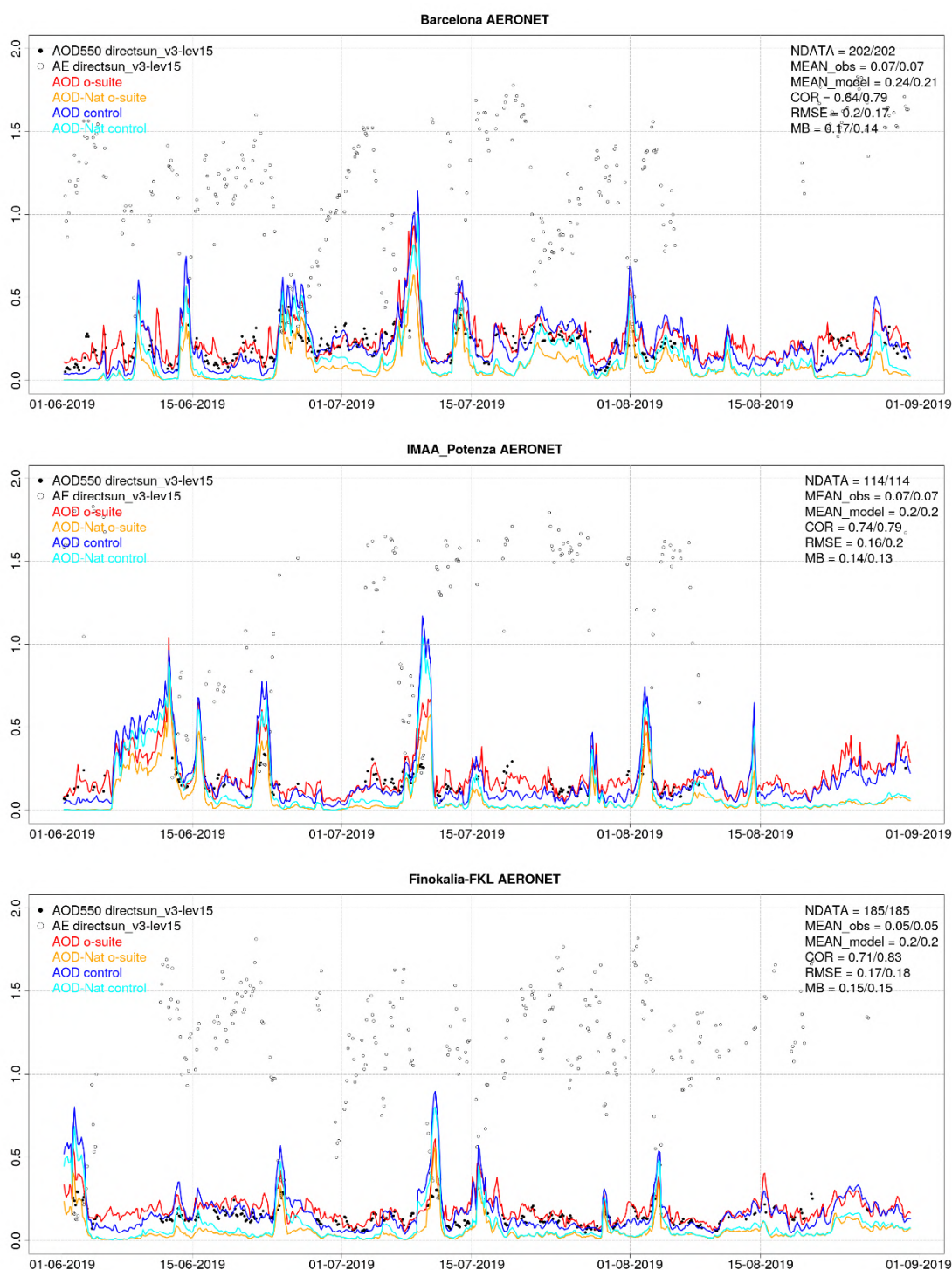
Figure 7.4.2: AOD from AERONET (black dot), AOD o-suite (red line), AOD control (blue line), AOD-Nat o-suite (orange line), AOD-Nat control (cyan line), for the study period over Barcelona (Spain), IMAA Potenza (Italy) and Finokalia (Crete). AOD-Nat corresponds to the natural aerosol optical depth that includes dust and sea-salt. Skill scores per each individual site and model (o—suite/control) are shown in the upper right corner (NDATA: available 3-hourly values used for the calculations, MEAN observations, MEAN model, COR, RMSE, MB).
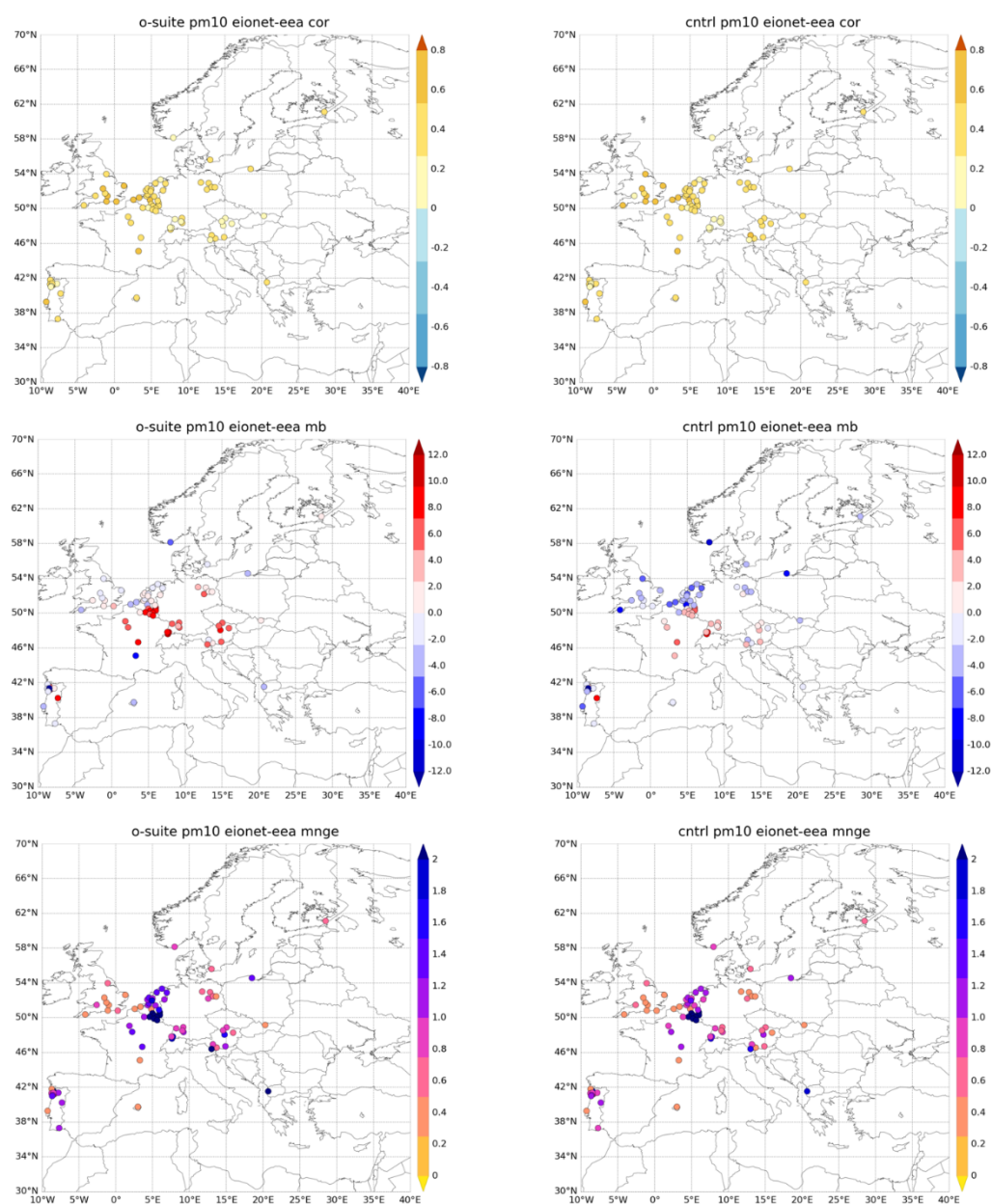
Figure 7.4.3: Skill scores (correlation coefficient, MB, RMSE and FGE) for 24-hour forecasts of CAMS o-suite and control for the study period. PM10 from EIONET are the reference. Only background suburban and rural available stations are displayed.
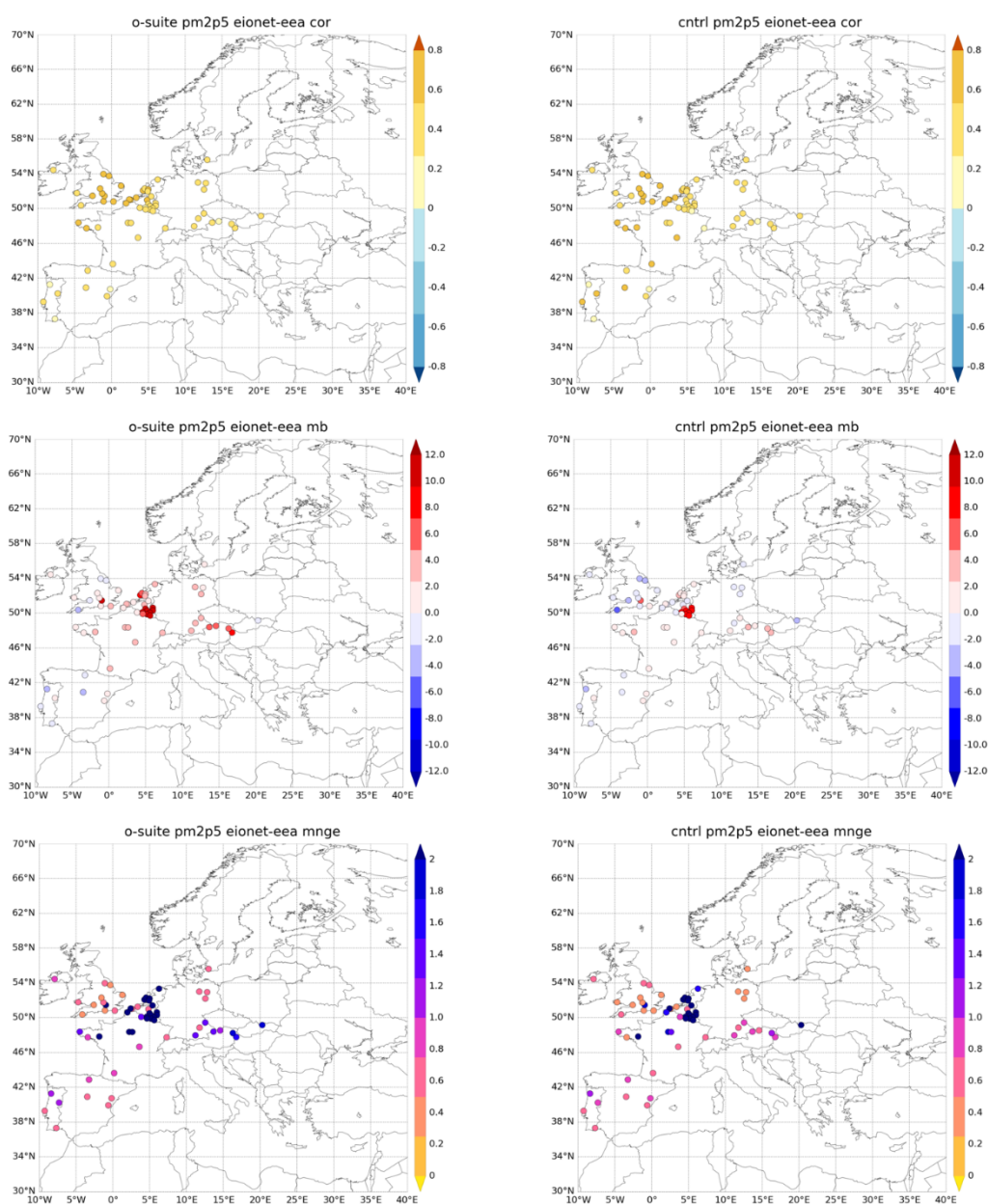
Figure 7.4.4: Skill scores (correlation coefficient, MB, RMSE and FGE) for 24-hour forecasts of CAMS o-suite and control for the study period. PM2.5 from EIONET are the reference. Only background suburban and rural available stations are displayed.
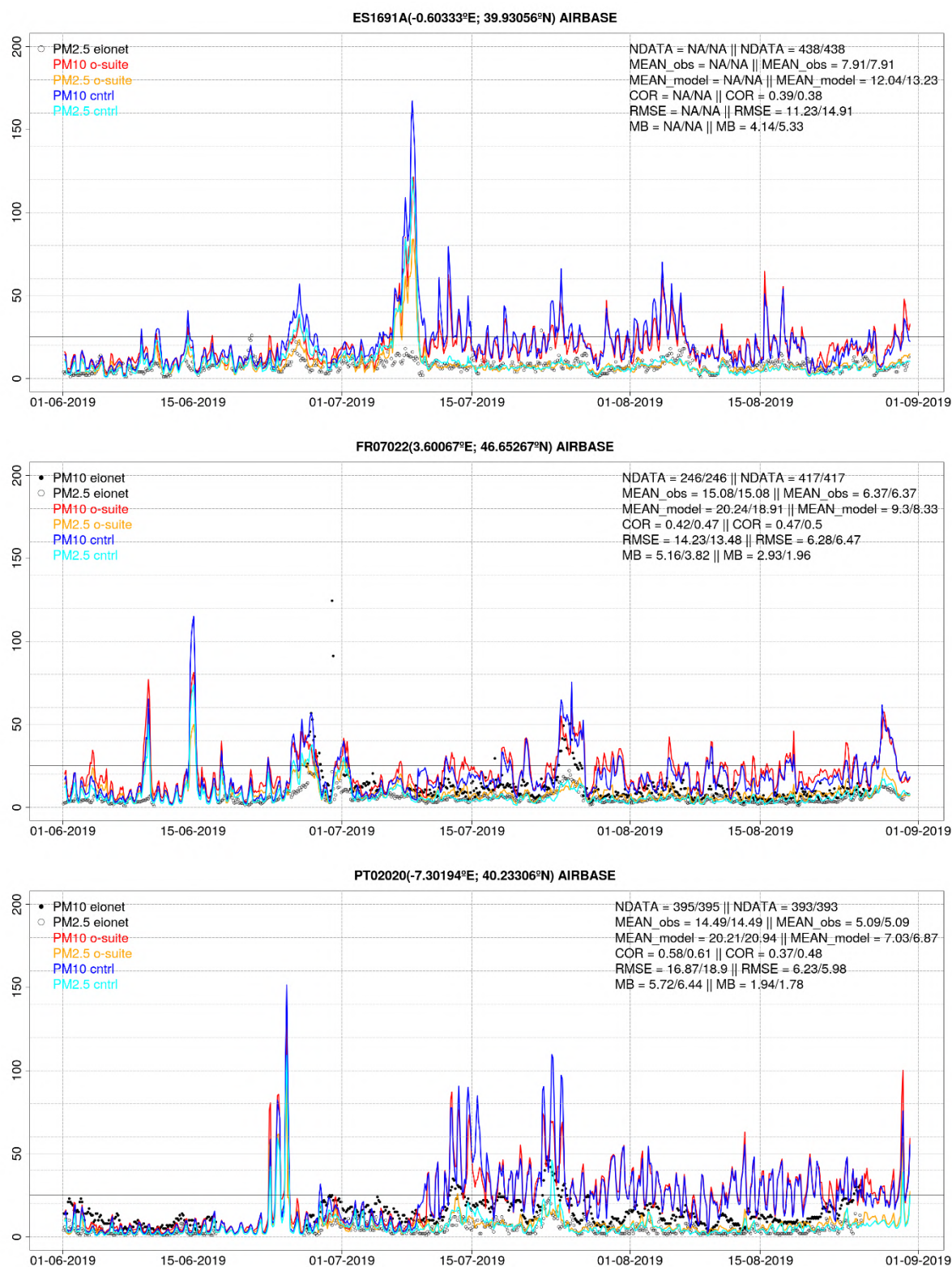
Figure 7.4.5: PM10 and PM2.5 Airbase observations (black and grey dots, respectively), PM10 and PM2.5 o-suite (red and orange lines, respectively) and PM10 and PM2.5 control (blue and cyan lines, respectively) for the study period over ES1691A (Spain), FR07022 (France) and PT02020 (Central inland Portugal).

8. Stratosphere

8.1 Validation against ozone sondes

In this section, we present the results of the stratospheric ozone evaluation against ozone soundings from the NDACC, WOUDC, NILU and SHADOZ databases. The ozonesondes have a precision of 3-5% (~10% in the troposphere for Brewer Mast) and an uncertainty of 5-10%. For further details see Cammas et al. (2009), Deshler et al. (2008) and Smit et al (2007). Model profiles of the o-suite are compared to balloon sondes measurement data of 44 stations for the period January 2013 to August 2019 (please note that towards the end of the validation period fewer soundings are available). As C-IFS-CB05 stratospheric composition products beyond O_3 in the o-suite is not useful we provide only a very limited evaluation of the control experiment. A description of the applied methodologies and a map with the sounding stations can be found in Eskes et al. (2018). The o-suite shows MNMBs within the range ±12%, for all regions and months (some exceptions with MNMBs of up to ±18% for single months in the high latitude regions). Figure 8.1.1. shows the results for the past year.

Fig. 8.1.2 compares the averaged profiles in each region during July 2019. The vertical distribution of stratospheric ozone is quite well represented for all regions by the o-suite, with little overestimation in all latitude bands (MNMBs between -2 to +6% for JJA 2019).

The control run shows a strong overestimation of stratospheric ozone in the upper stratosphere, and an underestimation between 50hPa and 300 hPa in the Antarctic and the Northern Midlatitudes, for the Arctic between 100 and 300 hPa.
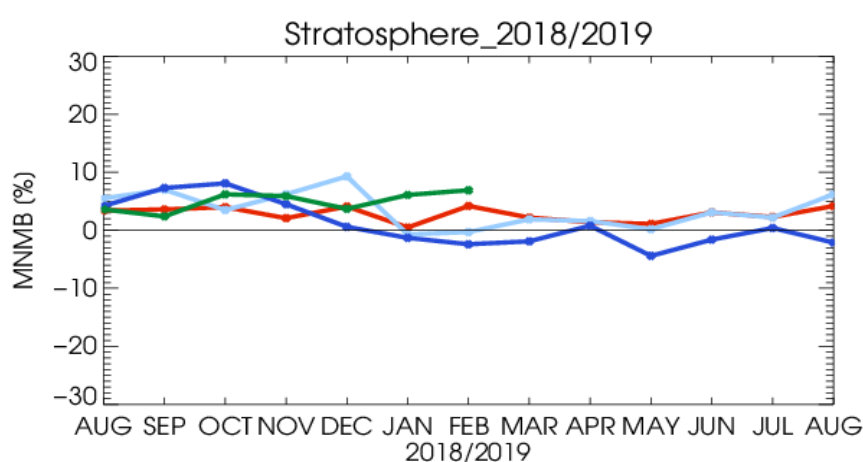
Figure 8.1.1: MNMBs (%) of ozone in the stratosphere from the o-suite against aggregated sonde data in the Arctic (light blue), Antarctic (dark blue) northern midlatitudes (red) and tropics (green). Period August 2018 to August 2019. The stratosphere is defined as the altitude region between 60 and 10 hPa in the tropics and between 90 and 10 hPa elsewhere.
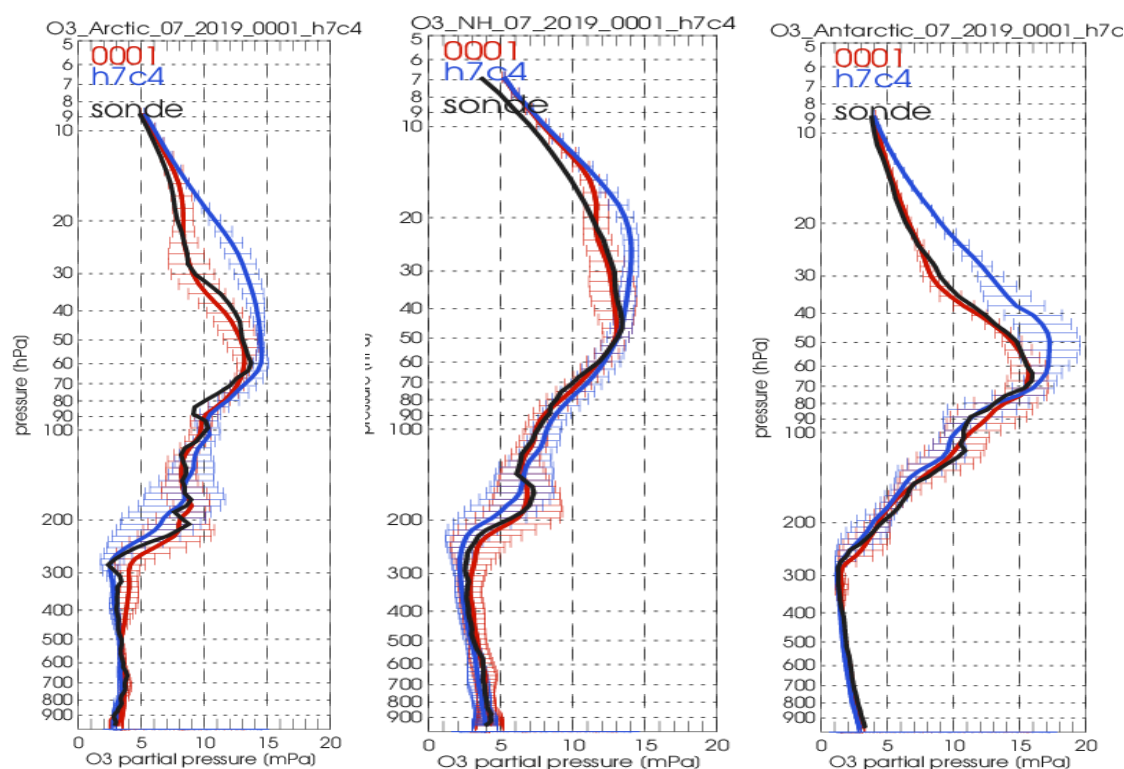
Figure 8.1.2: Comparison between mean O_3 profiles (units: mPa) of o-suite (red), and control (blue) in comparison with observed ozonesonde profiles (black) for July 2019 for the various latitude bands: Arctic, NH-mid latitudes and Antarctic.
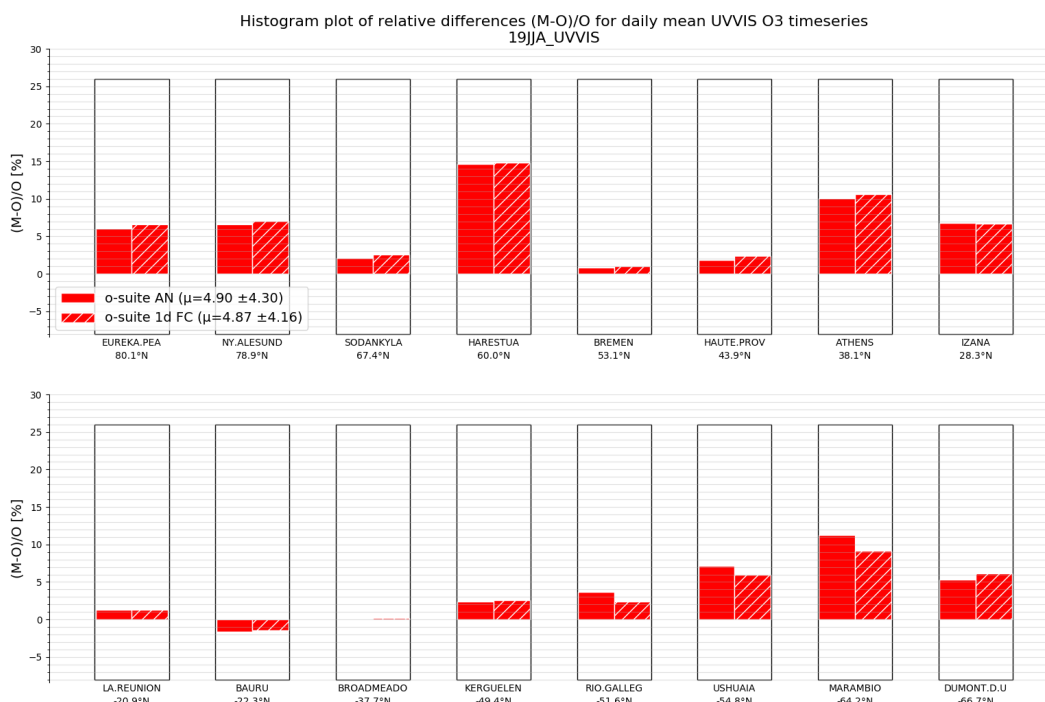
Figure 8.2.1. Relative biases during quarter JJA 2019 for 16 UVVIS stations measuring stratospheric ozone columns with ZENITH measurement geometry (stations sorted with decreasing latitude). The overall relative bias is positive for all latitudes and comparable to the typical measurement uncertainty of 5% for most sites.
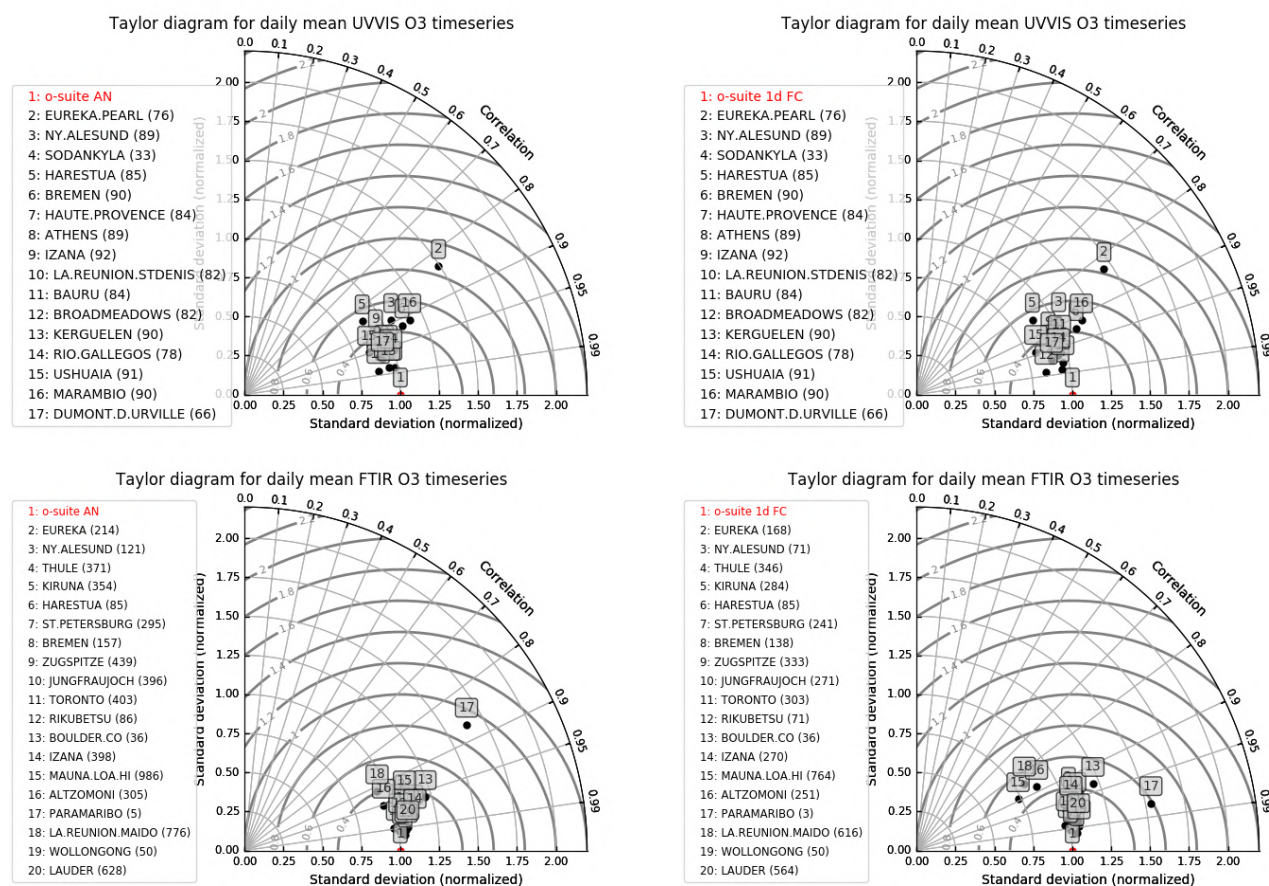
Figure 8.2.2. Taylor diagrams relating the standard deviations for the model and ground-based stratospheric column time series and their correlation for the time period JJA-2019. All time-series are normalized such that the std of the model is 1. The performance for the o-suite analysis is slightly better (averaged correlation is 0.97 for FTIR and 0.94 for UVVIS) compared to the 1-day forecasts (averaged correlation is 0.94 for FTIR and 0.93 for UVVIS). Again, the correlation for the tropical sites are worse in the 1-day forecasts compared to the analysis.

8.2 Validation against observations from the NDACC network

UVVIS and FTIR stratospheric columns

Since the start of the CAMS27 project, the number of UV-VIS Zenith ozone measurements have increased on NDACC. Currently fifteen sites provided data in the recent quarter allowing for a representative picture on the latitude dependence of the model data.

The systematic uncertainty of the UVVIS measurements is typically 5%, hence the relative biases for most sites for both the AN and 1d FC of the o-suite are very close to each other and within the uncertainty ranges, see Figure 8.2.1. The averaged bias for the 16 UVVIS sites is comparable to the measurement uncertainty of 5%, the averaged correlation is above 0.9.

The correlations between the sites and the model are presented in the Taylor diagrams in Figure 8.2.2. Again, the o-suite AN and 1d FC perform very similarly in correlation coefficients.
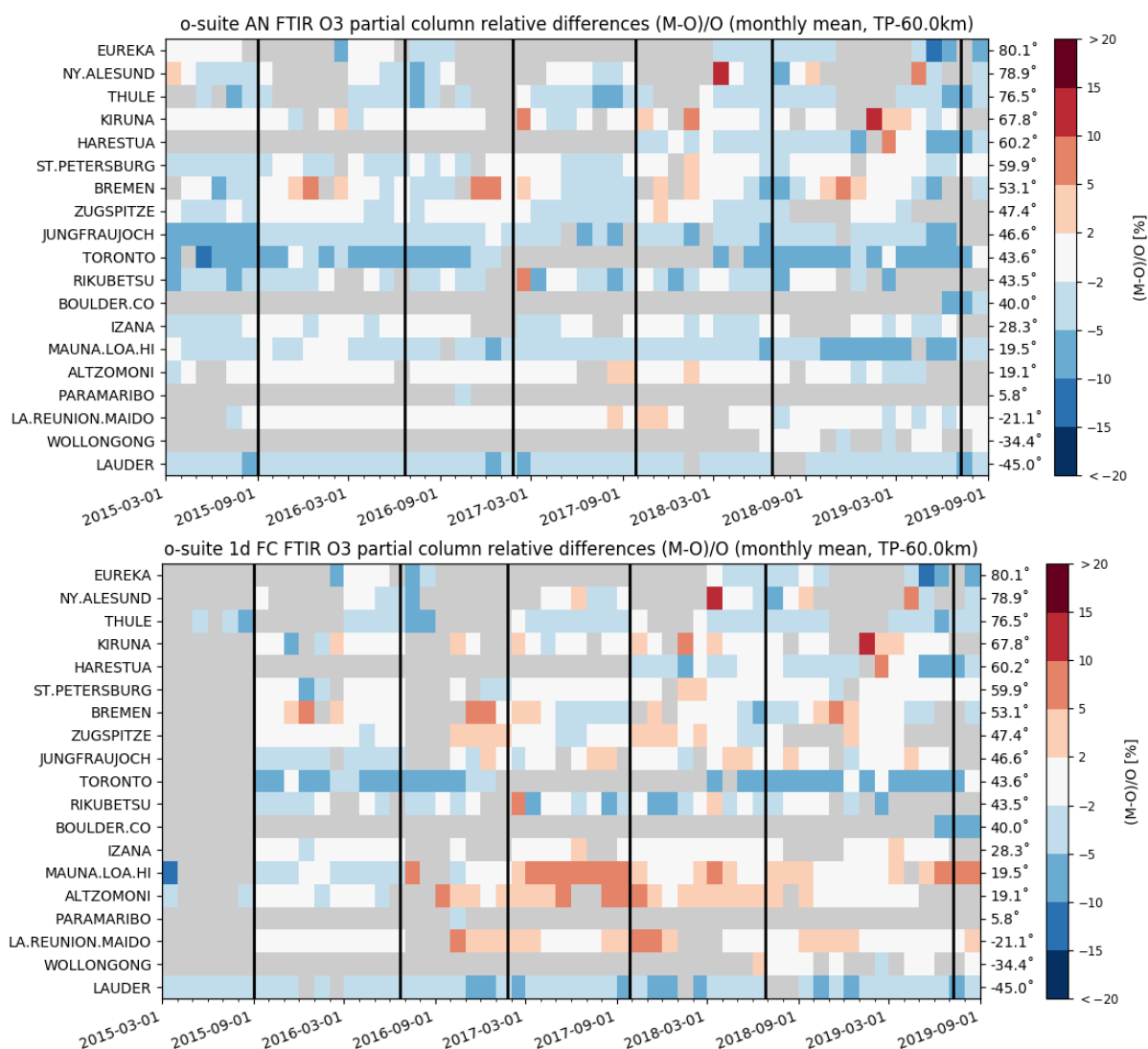
Figure 8.2.3 Time series of monthly mean relative differences for stratospheric FTIR columns along with model cycle updates (black vertical lines) (o-suite AN top, o-suite 1d FC bottom). The stratospheric FTIR columns for the tropical sites at Izana, Mauna Loa, Altzomoni and Reunion show a higher overestimation for the 1dFC compared to the AN.

Figure 8.2.3 depicts the FTIR stratospheric columns showing a discontinuity in the o-suite 1d FC model for the tropical sites (Mauna Loa, Altzomoni and Reunion) in the June 2016 model update. The worse performance of the tropical sites is also seen in lower correlations in Figure 8.2.2)

Profile comparison using LIDAR and MWR

In this section we present a comparison between the CAMS o-suite and control products against MWR and LIDAR observations from the NDACC network. A detailed description of the instruments and applied methodologies for all NDACC instruments can be found at http://nors.aeronomie.be. MWR (microwave) at Ny Alesund (79°N, 12°E, Arctic station) and Bern (47°N, 7°E, northern midlatitude station). LIDAR at Observatoire Haute Provence (OHP), France (43°N, 5.7°E, altitude 650m), Hohenpeissenberg, Germany (47°N, 11°E, altitude 1km) and Mauna Loa, Hawaii (19.5°N, 204°E, altitude 3.4km).

MWR.O3 VMR profile differences (M-O)/O
(weekly mean, 25.0km - 65.0km, o-suite AN, NY.ALESUND (lat.=78.9°), 2016-02-01 till 2019-08-31, 8726 meas.)

MWR.O3 VMR profile differences (M-O)/O
(weekly mean, 25.0km - 65.0km, o-suite AN, BERN (lat.=47.0°), 2015-03-01 till 2019-08-19, 12280 meas.)

LIDAR.O3 number density profile differences (M-O)/O
(weekly mean, 20.0km - 45.0km, o-suite AN, HOHENPEISSENBERG (lat.=47.8˚), 2015-03-06 till 2019-08-30, 513 meas.)

LIDAR.O3 number density profile differences (M-O)/O
(weekly mean, 20.0km - 45.0km, o-suite AN, HAUTE.PROVENCE (lat.=43.9˚), 2015-09-01 till 2019-03-29, 312 meas.)

LIDAR.O3 number density profile differences (M-O)/O
(weekly mean, 20.0km - 45.0km, o-suite AN, MAUNA.LOA.HI (lat.=19.5°), 2018-04-05 till 2019-08-29, 232 meas.)

Figure 8.2.4: Comparison of the weekly mean profile bias between the O_3 mixing ratios of o-suite and the NDACC station at Ny Alesund, Bern, Hohenpeissenberg, OHP and Mauna Loa. For the LIDAR stations, the measurement uncertainty above 35km is comparable to the observed profile bias.

At OHP, Hohenpeissenberg and Mauna Loa (LIDAR, see Figure 8.2.4), the o-suite slightly overestimates the observed ozone (<10%) between 25km and 35km. The uncertainty on the LIDAR concentration increases with altitude and above 35km the observed differences are comparable to the measurement uncertainty
(>10%, see http://nors.aeronomie.be/projectdir/PDF/NORS_D4.2_DUG.pdf).

8.3 Comparison with dedicated systems and with observations by limb-scanning satellites

This section compares the output of the o-suite for the last period with observations by limb-scanning satellite instruments, using the methodology described by Lefever et al. (2015). We also include the comparisons for the o-suite 4th day forecasts (96h to 120h) of stratospheric ozone. These forecasts are represented by dotted lines in the figures.

All datasets are averaged over all longitudes and over the three most interesting latitude bands for stratospheric ozone: Antarctic (90°S-60°S), Tropics (30°S-30°N) and Arctic (60°N-90°N). In order to provide global coverage, the two mid-latitude bands (60°S-90°S and 60°N-90°N) are also included in some comparisons with satellite observations.

The level-2 data from limb scanning instrument used in this section are:

- ACE-FTS version 3.6, on board SCISAT-1
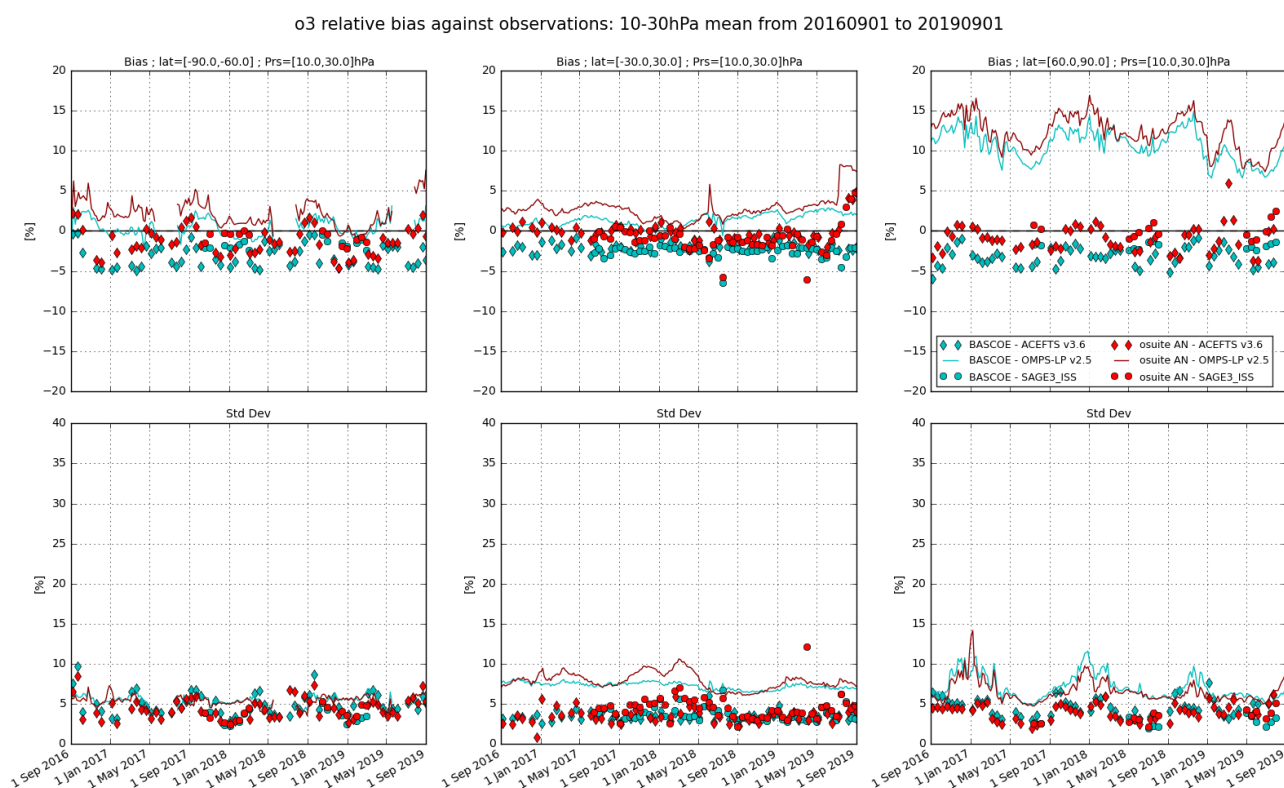
Figure 8.3.1: Time series comparing models to observations for the period 2016-09-01 to 2019-09-01 in the upper stratosphere (10-30hPa averages): o-suite analyses (red) and BASCOE (cyan) vs OMPS-LP (solid), ACE-FTS (diamonds) and SAGE-III (bullets). Top row: normalized mean bias (model-obs)/obs (%); bottom row: standard deviation of relative differences (%).

- SAGE-III version 5.1, on board the International Space Station (ISS); among the 3 different ozone profiles delivered by the solar occultation (denoted Mesospheric, MLR and AO3), we use the AO3 retrieval which is recommended by the mission science team.

- OMPS-LP version 2.5, on board NPP

For reference, we include also the BASCOE analyses which are very constrained by the AURA MLS offline profiles.

Figure 8.3.1 to 8.3.3 present, in the upper row, the timeseries over the last 36 months of the bias of the o-suite against the three satellite measurements for respectively three layers of the stratosphere (10-30hPa upper, 30-70hPa middle, and 70-100hPa lower and UTLS); the bottom row of the figures shows the standard deviation of the differences and can be used to evaluate the random error in the analyses.

In the tropics for the 70-100hPa region, the comparison with all instruments is unreliable (highly scattered bias and large standard deviations)
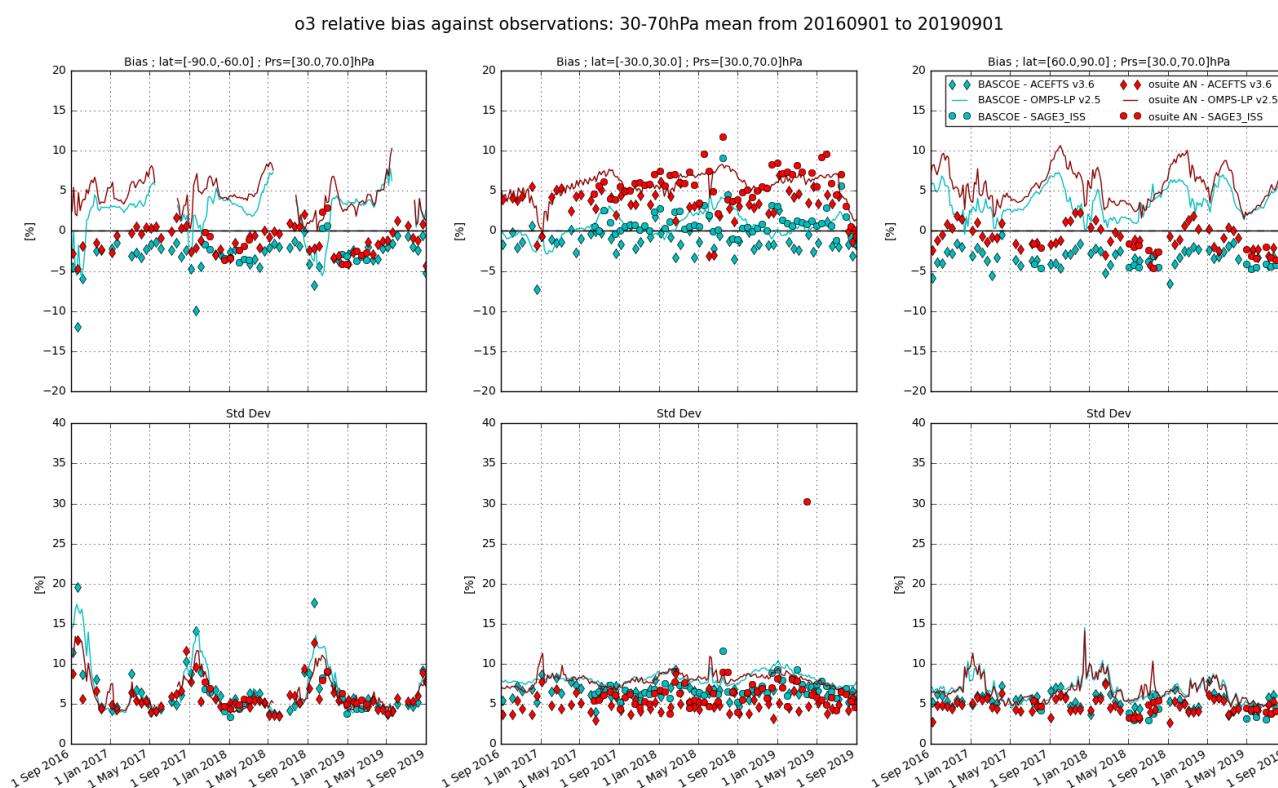
Figure 8.3.2: Time series comparing models to observations for the period 2016-09-01 to 2019-09-01 in the middle stratosphere (30-70hPa averages): o-suite analyses (red) and BASCOE (cyan) vs OMPS-LP (solid), ACE-FTS (diamonds) and SAGE-III (bullets). Top row: normalized mean bias (model-obs)/obs (%); bottom row: standard deviation of relative differences (%).

The agreement with ACE-FTS is good: the bias is generally within ±5% for all regions.

The SAGE-III onboard ISS provide observations since June 2017. The latitudinal coverage is more limited than ACE-FTS; the polar regions are not covered for long periods of time (e.g south polar region in 2019 only available in February). Where available, the agreement of the osuite with SAGE-III is good, with biases similar to those observed against ACE-FTS, except in the tropics in the 30-70hPa region where they are more positive (3-13%).

Compared to OMPS-LP, there is an almost systematic overestimation by the o-suite; the biases are more variable and more marked than for the other instruments (10% to 15% in the north polar at 10-30hPa region, up to 10% at 30-70hPa and up to 20% at 70-100hPa).

The bias of BASCOE against the satellite observations for the considered regions is systematically lower, but follows a similar evolution as the o-suite.

Figure 8.3.4 to 8.3.7 display vertical profiles of the relative biases between the o-suite or BASCOE and the satellite measurements. The difference is averaged over the most recent 3-month period considered in this validation report, i.e. June to August 2019.

Figure 8.3.3: Time series comparing models to observations for the period 2016-09-01 to 2019-09-01 in the lower stratosphere (70-100hPa averages): o-suite analyses (red) and BASCOE (cyan) vs OMPS-LP (solid), ACE-FTS (diamonds) and SAGE-III (bullets). Top row, normalized mean bias (model-obs)/obs (%);bottom row, standard deviation of relative differences (%).

The bias against each instrument remains within ±10% between 20km and 35km.

All o-suite profiles present a common feature of a slight overestimation at around 30km, followed by a stronger underestimation at around 40km, which is evidenced in the 4th day forecast.

At the higher part of the south polar profiles, an overestimation of ~1.5 ppm appears above 55km compared to ACE-FTS and 0.5 hPa compared to MLS.

It must be noted that the different instruments have a variety of spatial and temporal coverage: for a 3 month period and over the latitude bands considered, OMPS and Aura MLS provide daily data with more than 40000 valid profiles, while ACE-FTS provides around 700 profiles in the polar region and 200 profiles in the tropics, and SAGE-III around 800 profiles in each mid-latitude band and the tropics, and only around 200 profiles in the north polar region (and only close to 60°).
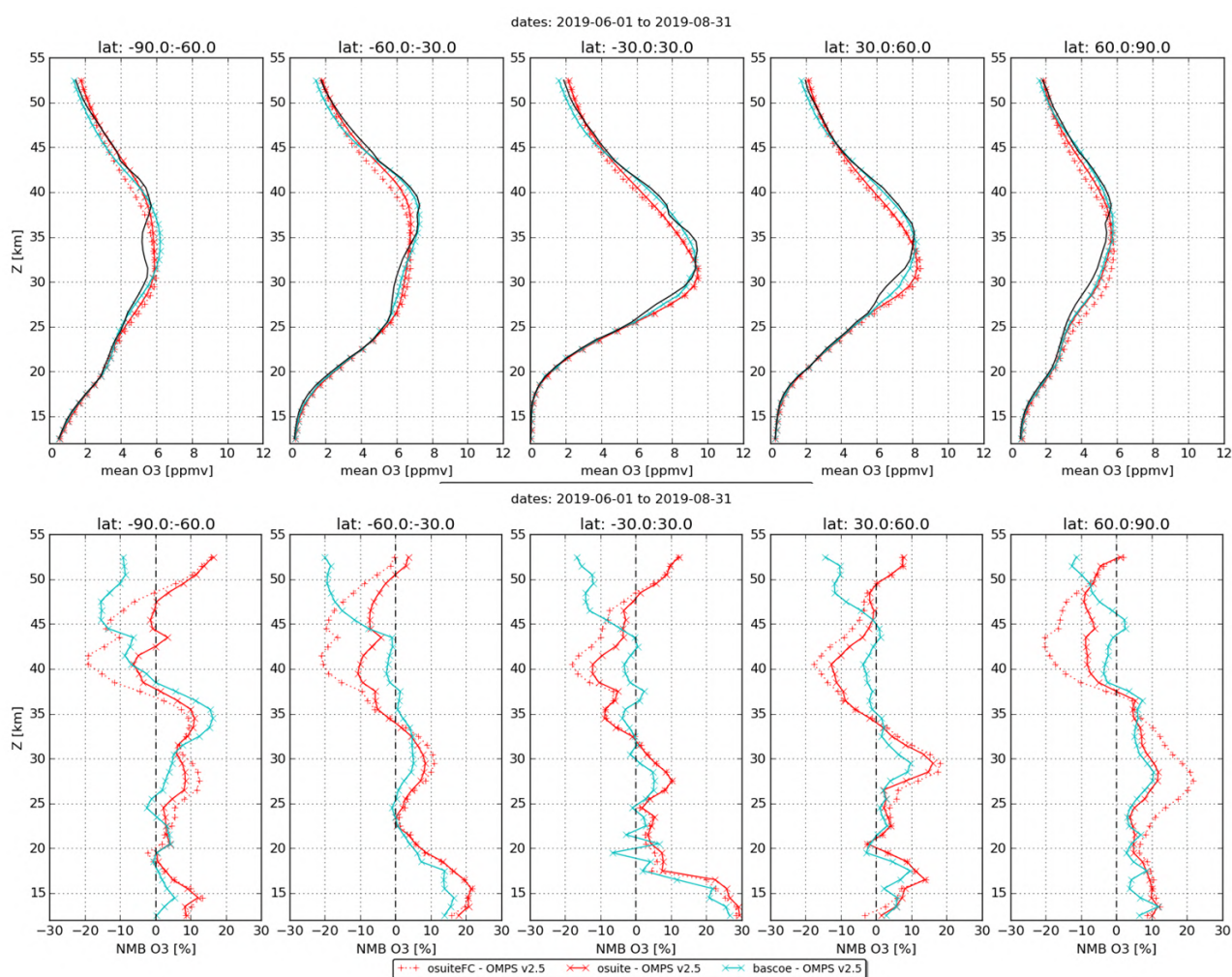
Figure 8.3.4: Mean value (top) and normalized mean bias (bottom) of the ozone profile between o-suite analyses (red, solid), o-suite forecasts 4th day (red, dotted) and BASCOE (cyan line) with OMPS-LP v2.5 observations for the period JJA 2019.
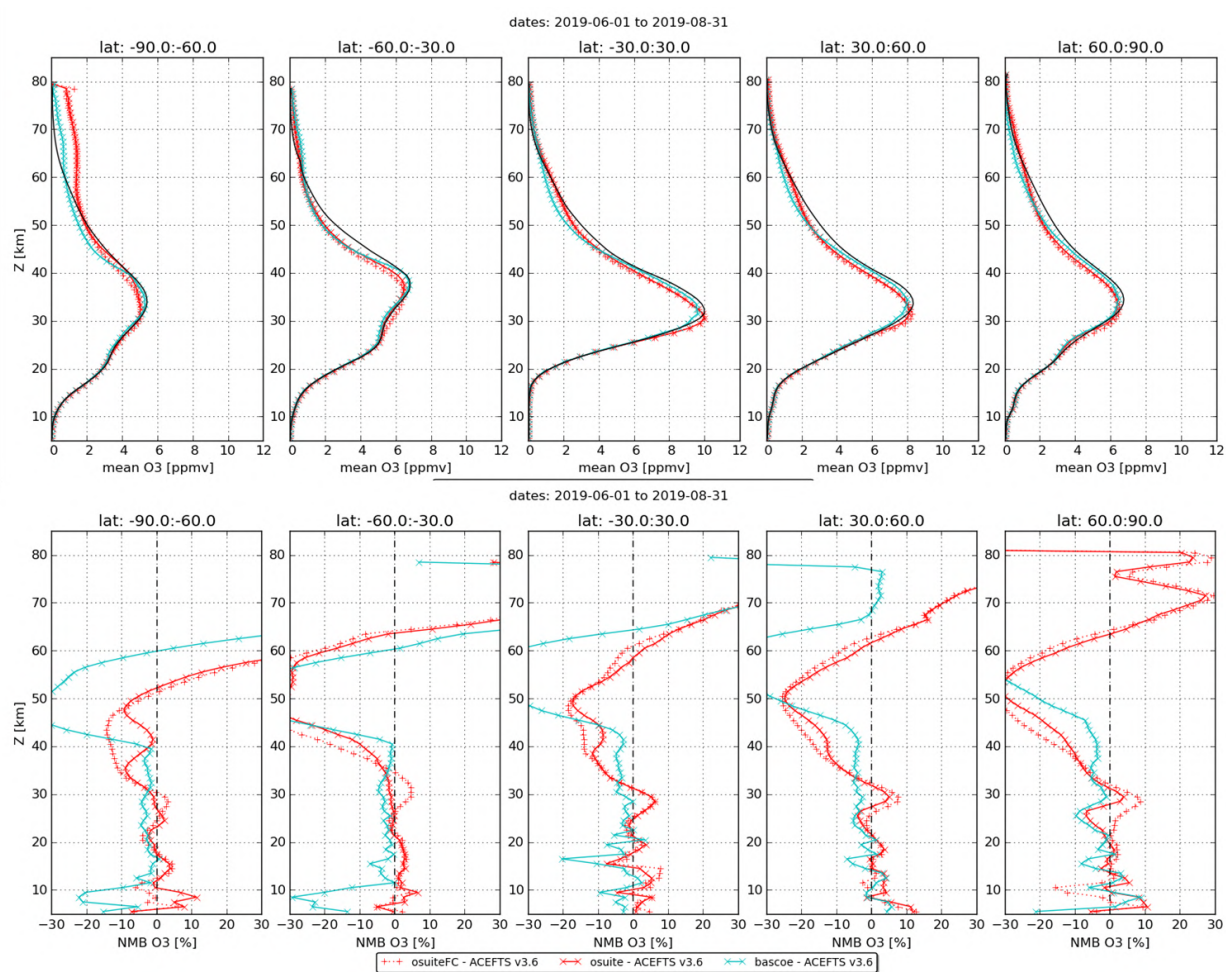
Figure 8.3.5: Mean value (top) and normalized mean bias (bottom) of the ozone profile between o-suite analyses (red, solid), o-suite forecasts 4th day (red, dotted) and BASCOE (cyan line) with ACE-FTS observations for the period JJA 2019.
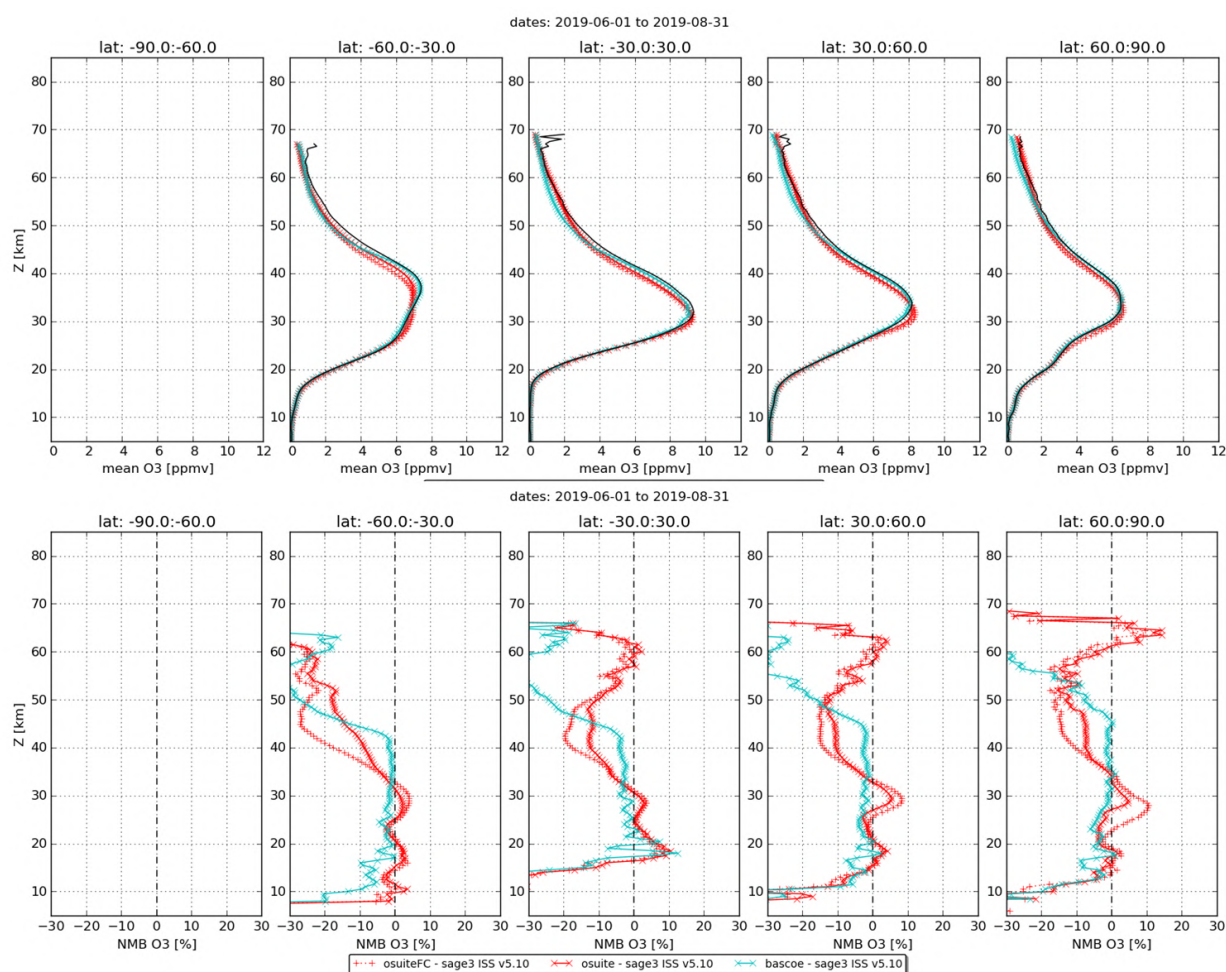
Figure 8.3.6: Mean value (top) and normalized mean bias (bottom) of the ozone profile between o-suite analyses (red, solid), o-suite forecasts 4th day (red, dotted) and BASCOE (cyan line) with SAGE-III observations for the period JJA 2019.
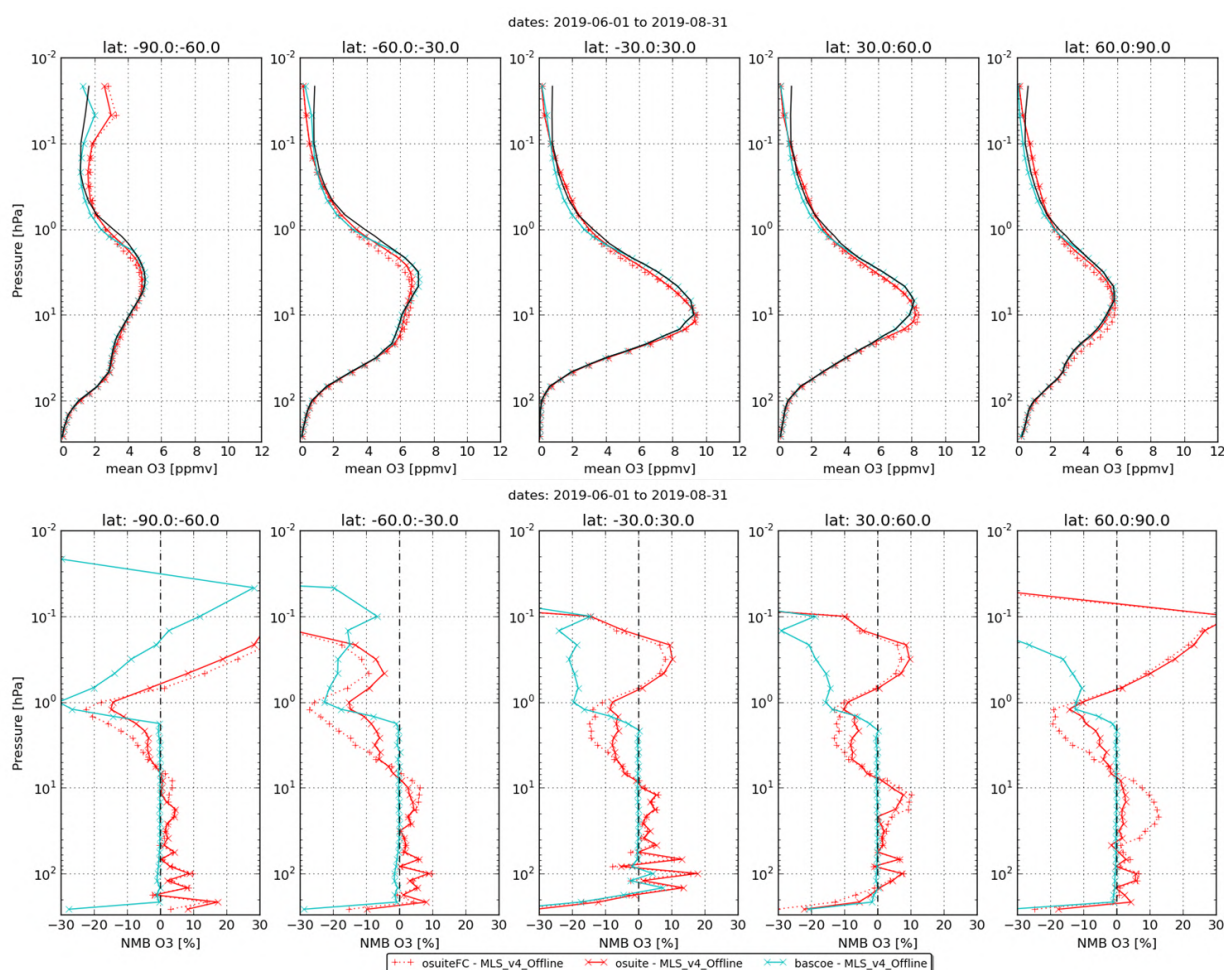
Figure 8.3.7: Mean value (top) and normalized mean bias (bottom) of the ozone profile between o-suite analyses (red, solid), o-suite forecasts 4th day (red, dotted) and BASCOE (cyan line) with MLS observations for the period JJA 2019.

8.4 Stratospheric NO₂

The CAMS model uses a tropospheric chemistry scheme in combination with a parameterization for stratospheric ozone. Stratospheric ozone is also well constrained by satellite observations. Therefore, the only useful product in the stratosphere is ozone, and all other compounds, including NO_2, should not be used, as demonstrated by the validation results presented here.

In this section, nitrogen dioxide from SCIAMACHY/Envisat satellite retrievals (IUP-UB v0.7) and GOME-2/MetOp-A satellite retrievals (IUP-UB v1.0) are compared to modelled stratospheric NO_2 columns. Monthly mean stratospheric NO_2 columns from SCIAMACHY and GOME-2 have relatively small errors on the order of 20% in the tropics and in mid-latitudes in summer and even lower errors at mid-latitudes in winter. As the time resolution of the saved model files is rather coarse and NO_x photochemistry in the stratosphere has a large impact on the NO_2 columns at low sun, some uncertainty is introduced by the time interpolation at high latitudes in winter.

As shown in Figure 8.4.1, amplitude and seasonality of satellite stratospheric NO_2 columns are poorly modelled with CB05-based chemistry runs including the more recent versions of the o-suite. The significant differences between observations and CB05 chemistry runs, i.e. a strong underestimation of satellite retrievals by models, can be explained by the missing stratospheric chemistry for these model versions. The only constraint on stratospheric NO_x is implicitly made by fixing the HNO_3/O_3 ratio at the 10 hPa level. This assumption, in combination with the changing model settings for stratospheric O_3 for control compared to MACC_CIFS_TM5, may explain some of the jumps we see in stratospheric NO_2. In any of these runs the stratospheric NO_2 is poorly constrained. It clearly indicates that stratospheric NO_2 in the latest versions of the o-suite is not a useful product and should be disregarded. However, model simulated values increased with an upgrade of the o-suite in February 2017, so that simulations are closer to the satellite observations for 2017, especially for northern hemisphere latitude bands where seasonality seems to have been reproduced (in contrast to the Southern Hemisphere) by the o-suite apart from the pronounced underestimation. O-suite values are larger than the control in 2017 at all latitude bands. The better agreement found for 2017 did not continue in the beginning of 2018 and values decreased to the magnitude of 2015-2016 runs at all latitude bands, followed however by a similar increase as in 2017 in late 2018 (development to be seen with the next reports).

Comparison of the o-suite from July 2012 until August 2014 with the other model runs and satellite observations shows that the previous version of the o-suite stratospheric NO_2 columns had a systematic low bias relative to those from MACC_fcnrt_MOZ and satellite observations for all latitude bands. For example, o-suite values are a factor of 2 smaller than satellite values between 60°S to 90°S for October 2013. Best performance was achieved with the MOZART chemistry experiments without data assimilation (MACC_fcnrt_MOZ, running until September 2014), especially northwards of 30°S. Details on the NO_2 evaluation can be found at: http://www.doas-bremen.de/macc/macc_veri_iup_home.html.
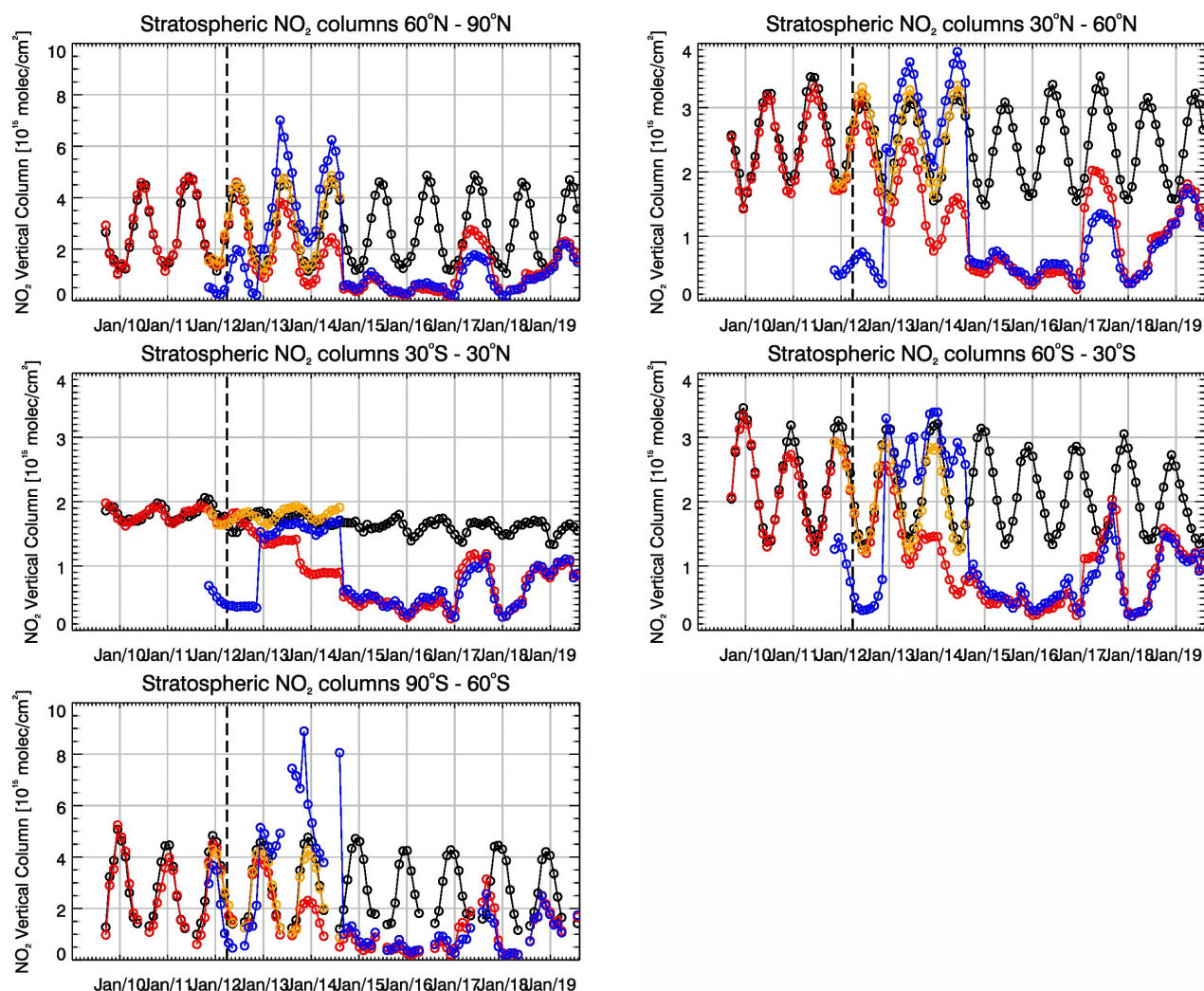
Figure 8.4.1: Time series of average stratospheric NO2 columns [10^{15} molec cm-2] from SCIAMACHY (up to March 2012) and GOME-2 (from April 2012, black) compared to model results (red: o-suite, blue: MACC fcnrt TM5/MACC CIFS TM5/control, orange: MACC fcnrt MOZ) for different latitude bands. See text for details. The blue line shows MACC_fcnrt_TM5 from November 2011 to November 2012, MACC_CIFS_TM5 results from December 2012 until August 2014 and control results from September 2014 onwards (the model run without data assimilation is termed control since Sep 2014). The vertical dashed black lines mark the change from SCIAMACHY to GOME-2 based comparisons in April 2012.

9. Validation results for greenhouse gases

This section describes the NRT validation of the pre-operational, high resolution forecast of CO_2 and CH_4 from 1st June 2018 to 1st June 2019 based on observations from 17 surface stations, located in Western Europe; 15 TCCON stations measuring XCO_2 and XCH_4 total columns, and 13 NDACC stations measuring partial and total CH_4 columns. We compare the observations to the high-resolution forecast experiments (*gqpe/gznv, Tco1279L137; 9x9 km*), coupled to the analysis experiment (*gqiq/gwx3, Tco399L137, 25x25 km*). The *gqpe* forecast experiment is using the IFS model cycle CY43R1 and has been officially implemented on 1st Nov. 2017. The *gznv/gwx3* experiments, based on IFS CY45R1, are used from 1st December 2018 on. This new experiment benefits from a couple of bugs fixed in the modelled biogenic fluxes which should result in an improved seasonal cycle in the northern hemisphere and some degradation in the tropics.

9.1 CH4 and CO2 validation against ICOS observations

The CO_2 and CH_4 simulations from the analysis and high-resolution forecast have been compared to the 17 ICOS stations. The near–real time data processing of the in-situ measurements is ensured by the Atmospheric Thematic Center (Hazan et al., 2016). Among the 17 stations we can distinguish three sites located on top of mountains (PUY, JFJ, CMN), two background sites (PAL, ZEP) and 12 tall towers. For the later we consider only in this report the highest sampling level which is at least at 100m above the ground.

For CO_2 the correlation coefficients are higher than 0.8 for all sites except one at Ispra, located in the Po valley, and poorly represented by the model due to the complex orography (Figure 9.1.1). The synoptic scale variability is overestimated at all stations except two (CMN, TOH). The best correlation coefficients are obtained for the background stations PAL, ZEP and JFJ. For CH_4, there is not much difference in the correlation coefficients for the background sites and the tall towers, which are close to 0.8 with two exceptions at Monte Cimone and Ispra where correlations are below 0.5. Overall, we notice a small degradation of the CH_4 correlations with the high-resolution forecast experiment compared to the analysis (Figure 9.1.1). This is particularly true at the Trainou tall tower. In this case several spikes are wrongly simulated by the models in summer, and their amplitude is getting worse with the high-resolution forecast. This is probably due to the vicinity of hot spot emissions in Paris area, either mislocated or overestimated in the emission inventory, whose influence is amplified when using higher resolution.

The figure 9.1.2 shows the time varying biases (observations minus model), averaged on a weekly basis, for all ICOS stations. We do not observe strong differences between the high-resolution forecast and the analysis simulations. The CO_2 biases display clear seasonal cycles (±10 ppm) at most sites with negative biases in Summer/Autumn, and positive biases in Winter/Spring. One example is detailed on Figure 9.1.3 for the Norunda tall tower located in Sweden. The negative biases in the autumn 2018 were significantly reduced compared to 2017, but it seems to be the reversed for the positive bias which have been increased at many sites in 2019 compared to 2018. This increase of the positive bias is more clearly seen in figure 9.1.5a showing the seasonal pattern of the model/observation differences averaged over the 17 sites. The maximum European bias
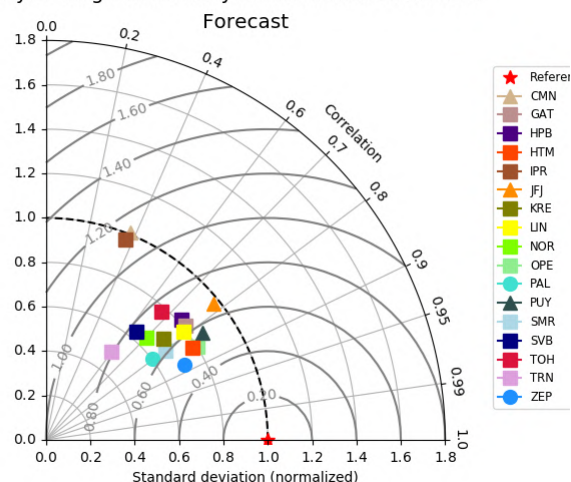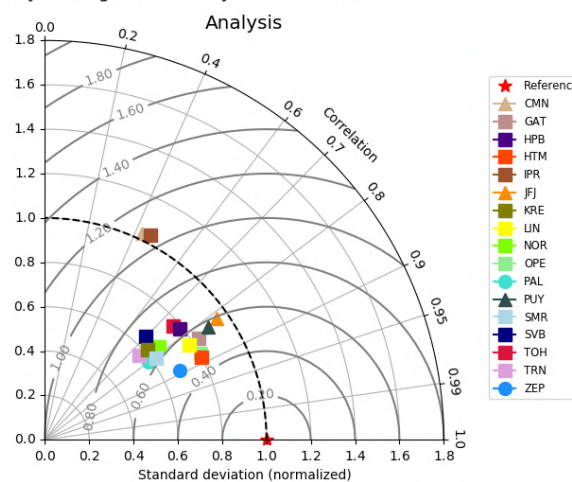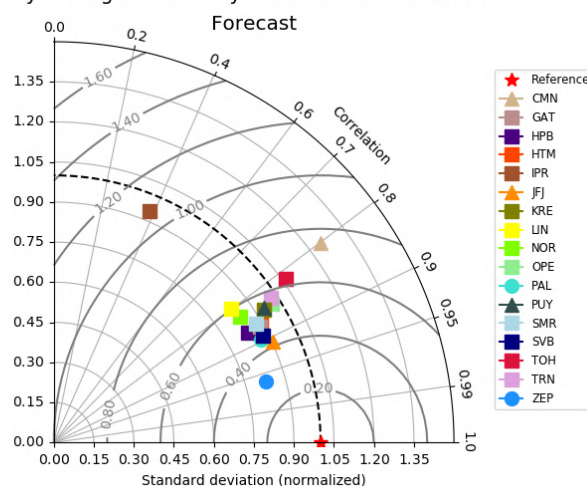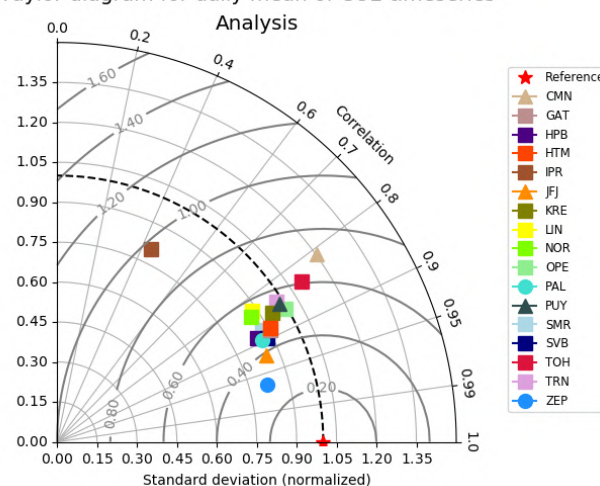
Figure 9.1.1: Taylor diagrams relating the standard deviations for the model and time series of CH$_4$ (above) and CO$_2$ (below) mole fractions and their correlation. The normalized standard deviation is calculated as the ratio observed SD / modelled SD (SD values lower than one mean an overestimation of the variability by the model). The left panels show the high-resolution forecast, and the right panels the analysis.

increases from 4 to 6 ppm from 2018 to 2019, and contrary to 2018 it does not start decreasing in June and remains high until at least August 2019 (end of the comparison). This feature can be observed for example at the Puy de Dôme site (Figure 9.1.4).

For CH$_4$ we observe a significant latitudinal gradient in the biases. At the Scandinavian sites the model overestimates (up to 50 ppb) the observations all year long. The example of Norunda (Figure 9.1.3) clearly shows that the CH$_4$ spikes are systematically too high in the model runs, which could indicate that the wetland emissions are overestimated. For sites located at lower latitudes, the bias is generally lower, and is getting negative in Summer/Autumn. It is even systematically negative at the mountain sites PUY (Figure 9.1.4) and JFJ. The latitudinal distribution of the CH$_4$ biases in Europe is clearly seen from the map in figure 9.1.5, despite two outliers at Ispra (Italy) and Trainou (France) showing positive biases due to the difficulty to represent the local topography (Ispra) and the Paris emissions (Trainou) as explained earlier.
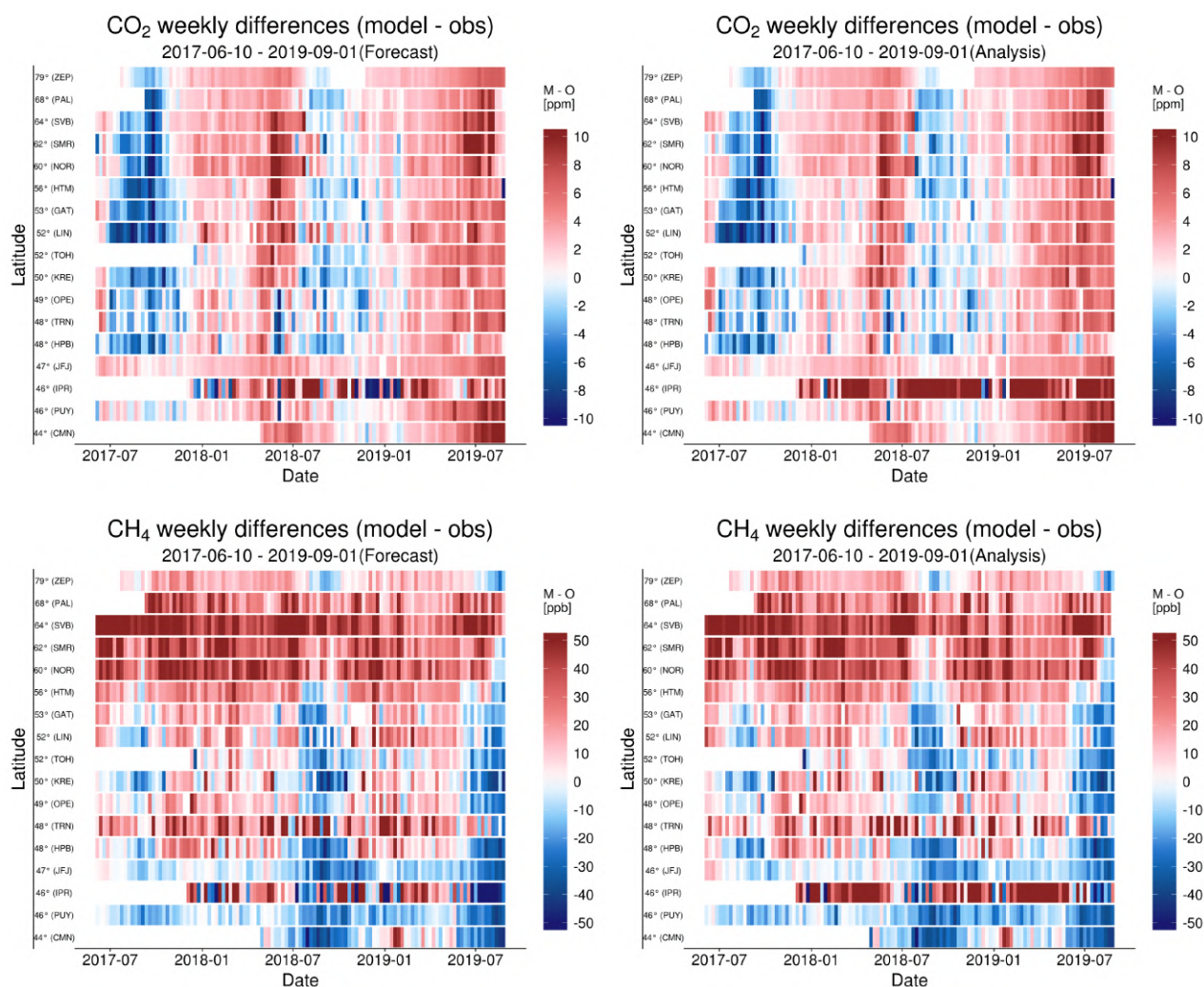
Figure 9.1.2: Mosaic plot of CO_2 (above, in ppm) and CH_4 (below, in ppb) biases of the CAMS products (left: high resolution forecast, right: analysis run) compared to surface station observations. Each vertical colored line represents a weekly mean.

We also do an evaluation of the model performance based on a composite of all European stations (Figure 9.1.6). Overall this composite comparison highlights the seasonal cycle of the CO_2 bias, and especially the recent increase of the bias, with a maximal offset up to 6 ppm in Summer 2019 (July-August). For several sites, like Norunda (Figure 9.1.3), a high bias is also observed in Spring 2018 resulting from the heatwave event in Northern Europe. It should also be noted that the coefficient correlations are significantly lower and RMS higher from April to September, during the growing season of the vegetation when the biospheric fluxes are maximum. For CH_4 we observe a mean positive bias up 20 ppb for most of the year, except in summertime (mean negative bias down to -10 ppb). However, considering the high latitudinal gradient (Figures 9.1.2 and 9.1.5), the mean CH_4 European bias must be taken with extreme caution.
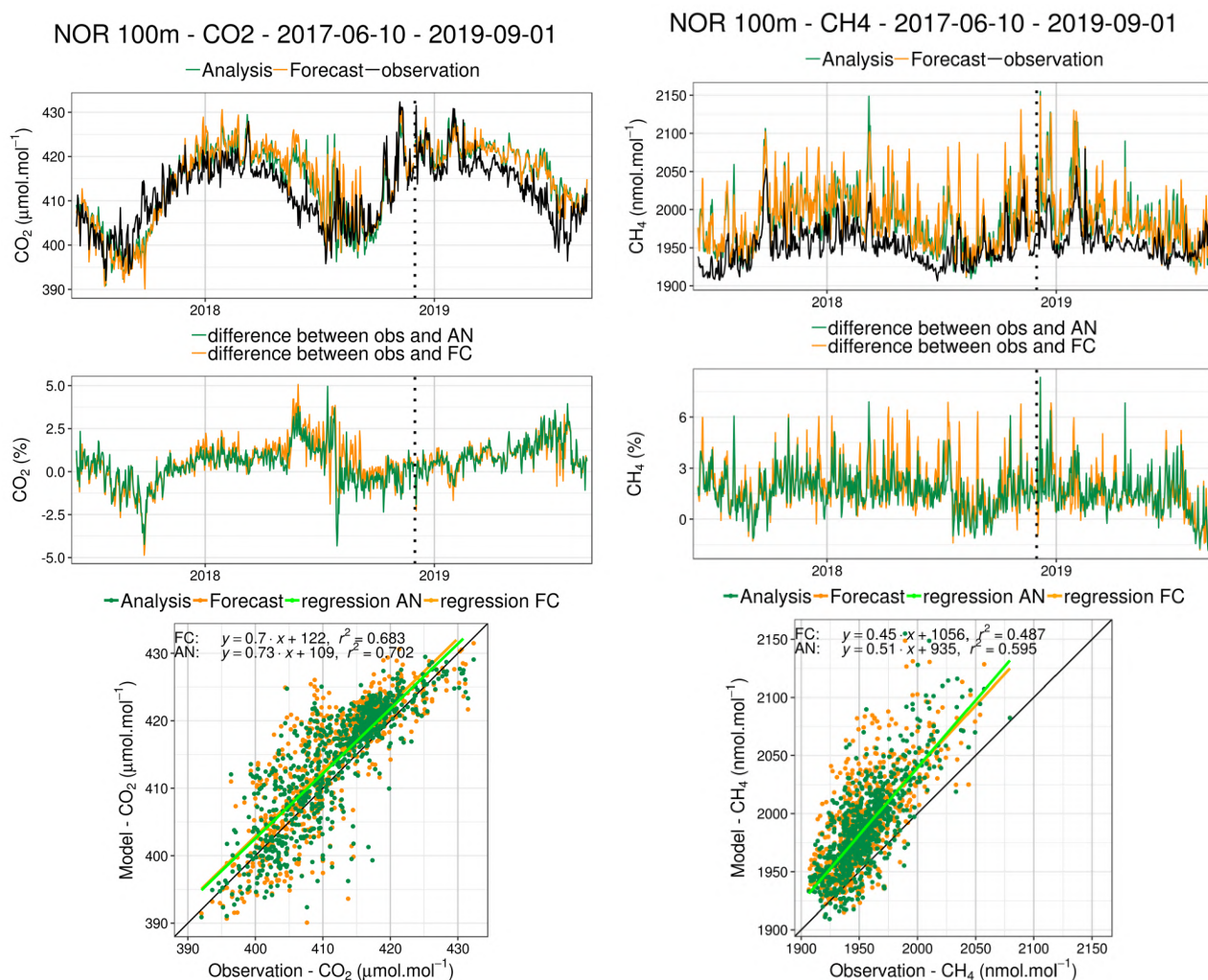
Figure 9.1.3: Comparison of CO_2 (left) and CH_4 (right) daily means observed (black) with the analysis run (green) and the high-resolution forecast (orange) at the Norunda tall tower. Middle: differences of the observations minus the simulations. Below: Linear fit between observations and simulations. The dashed vertical line represents the change of experiments in December 2018.
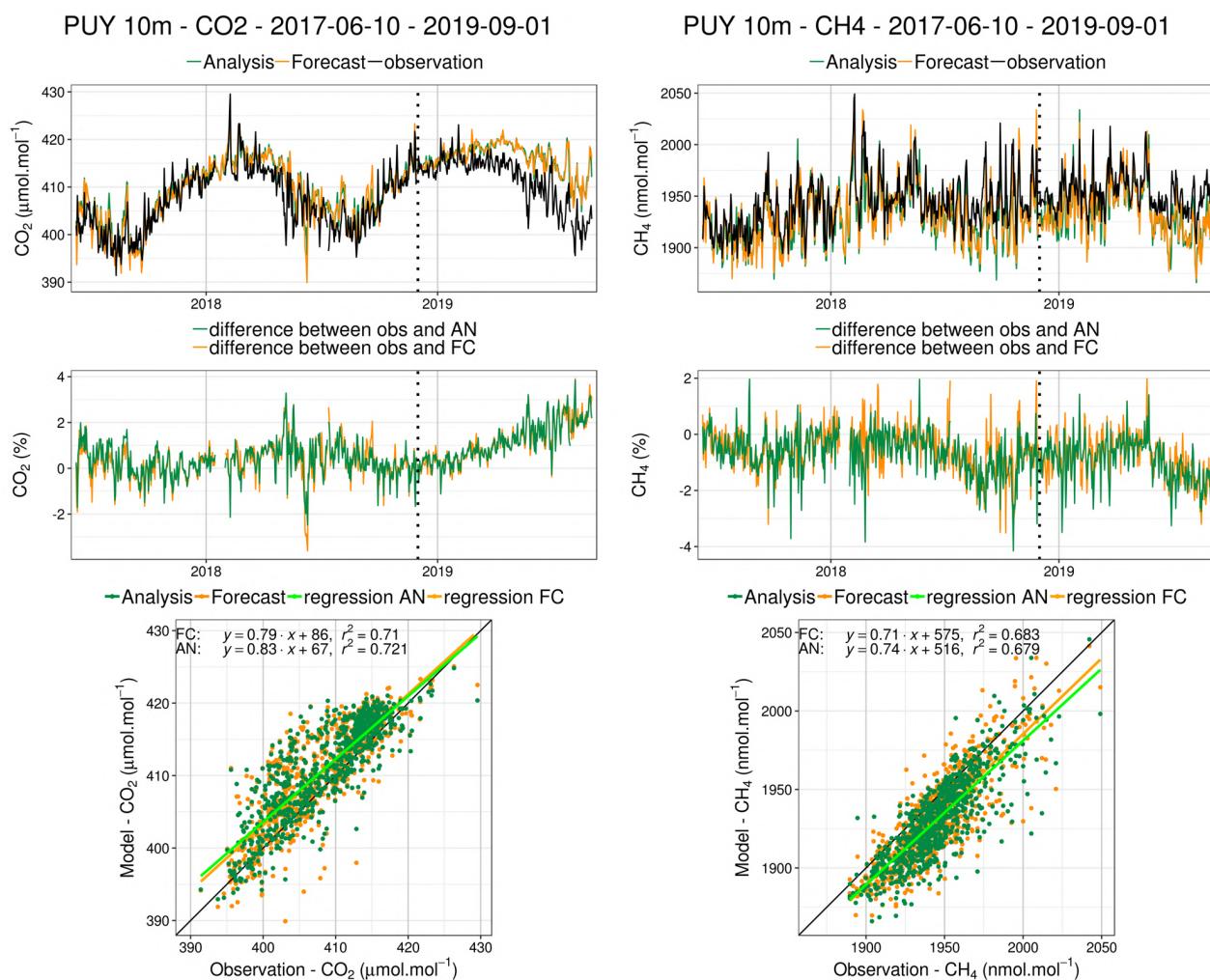
Figure 9.1.4: Same as figure 9.1.3 for Puy de Dôme site.
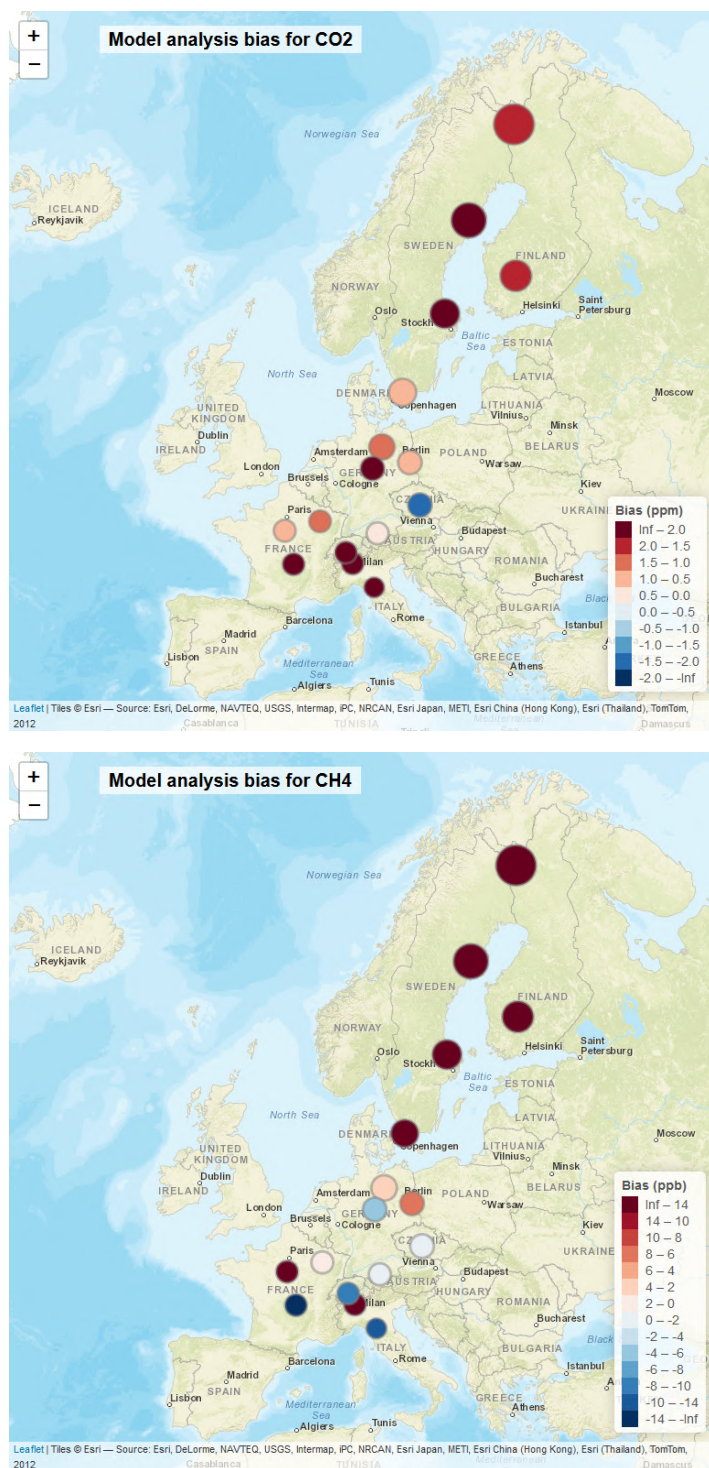
Figure 9.1.5: Mean biases (model – observations) for CO_2 (top, in ppm) and CH_4 (bottom, in ppb) at ICOS stations. The size of the points is not informative.

Analysis - Monthly metrics in the northern hemisphere
CMN-GAT-HPB-HTM-IPR-JFJ-KRE-LIN-NOR-OPE-PAL-PUY-SMR-SVB-TOH-TRN-ZEP
CO_2 (μmol.mol^{-1})
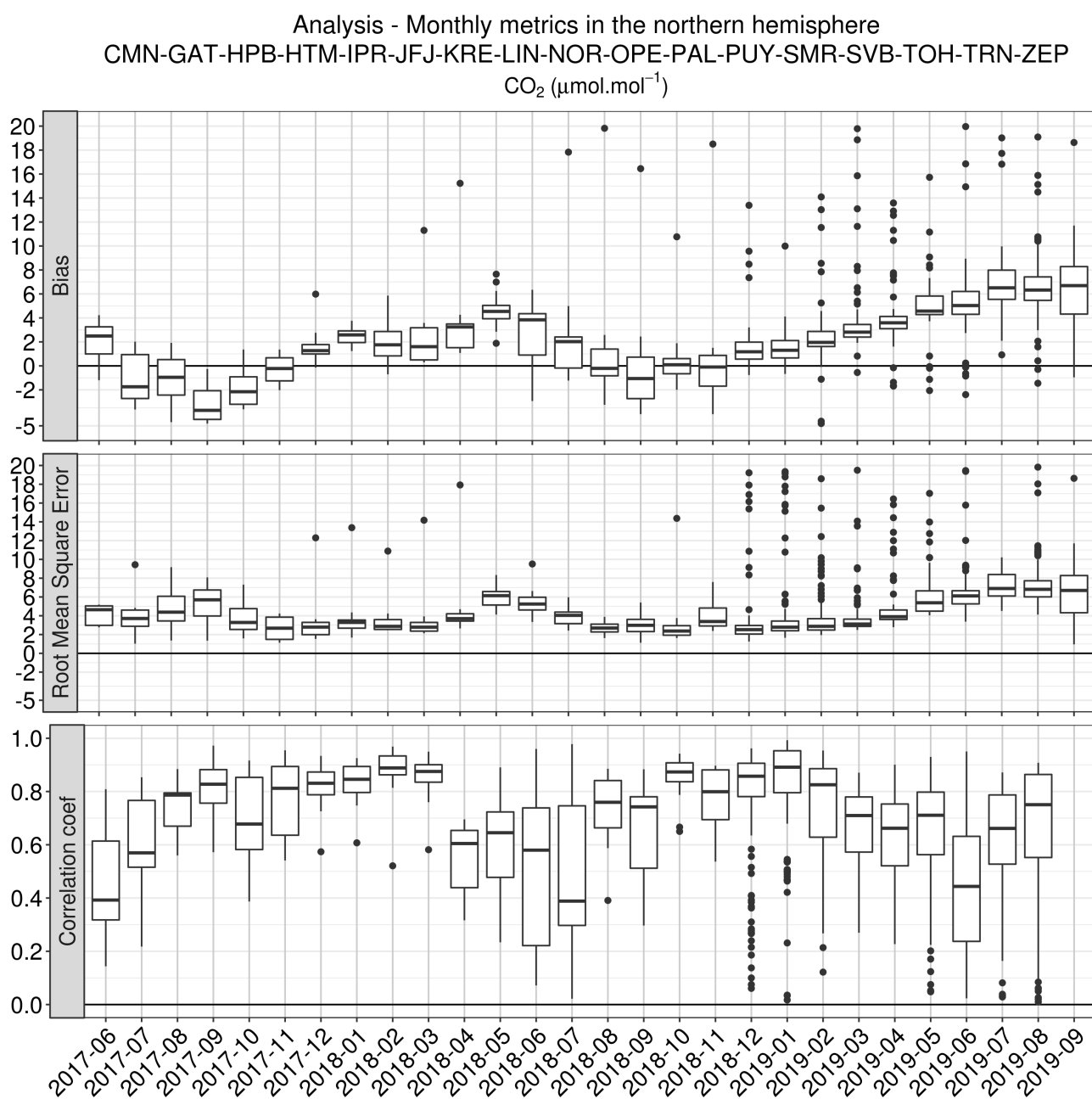
Figure 9.1.6a: Monthly statistics (bias, RMSE, correlation coefficients) of the analysis experiment compared to CO_2 surface measurements at ICOS sites. The results obtained for all European sites (see the list of sites in the title) are averaged. September 2019 is not representative, since using only one day.

Analysis - Monthly metrics in the northern hemisphere
CMN-GAT-HPB-HTM-IPR-JFJ-KRE-LIN-NOR-OPE-PAL-PUY-SMR-SVB-TOH-TRN-ZEP
CH_4 (nmol.mol^{-1})
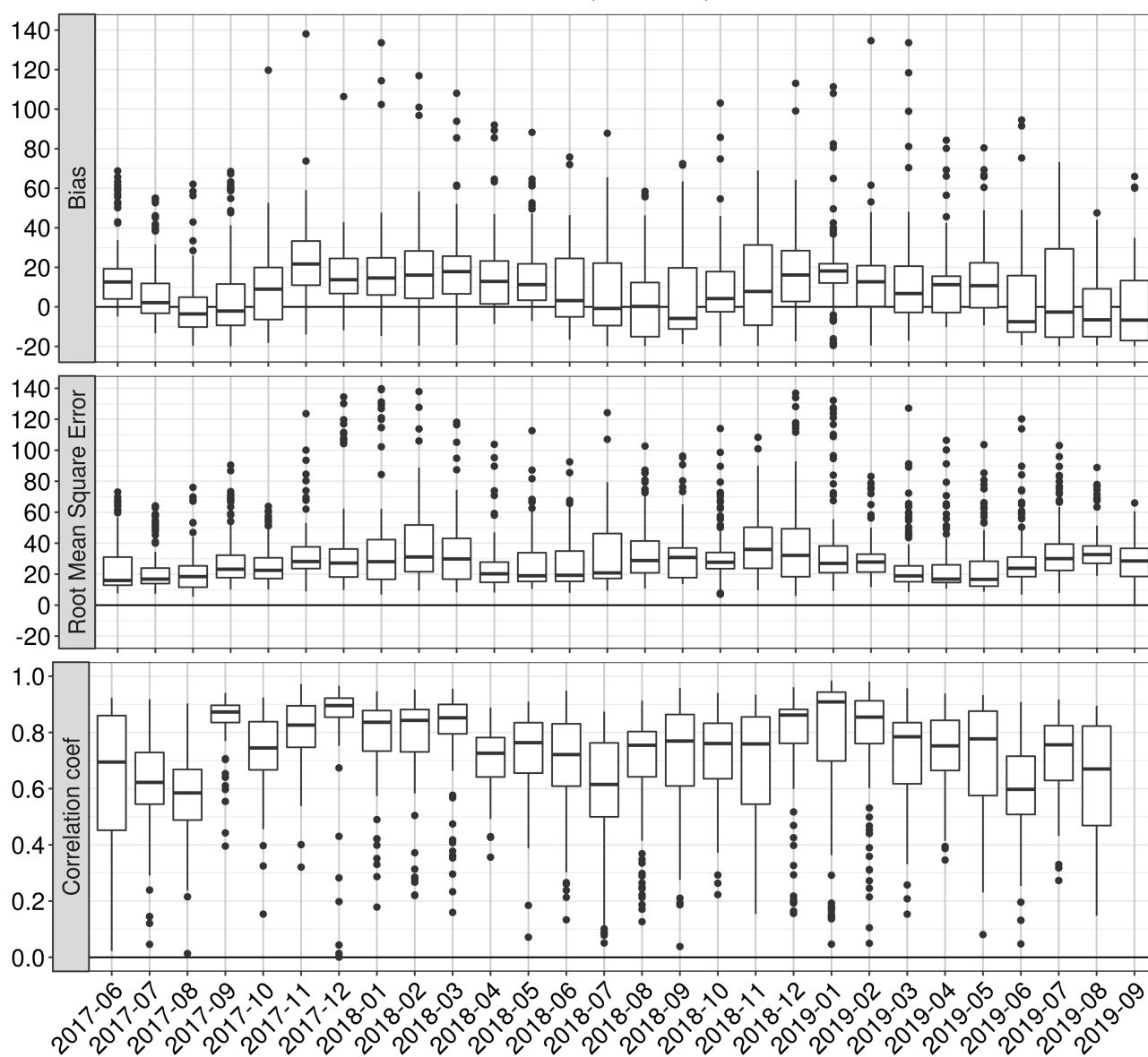
Figure 9.1.6b: Same as Figure 9.1.3a for CH_4.

9.2 CH$_4$ and CO$_2$ validation against TCCON observations

For the validation column averaged mole fractions of CO$_2$ and CH$_4$ (denoted as XCO$_2$ and XCH$_4$) from the Total Carbon Column Observing Network (TCCON) are used. Column averaged mole fractions provide different information than the in-situ measurements and are therefore complementary to the in-situ data.

The validation routines used for TCCON data are the same as used for the NDACC network and are documented in Langerock et al. (2015). The routines have been adapted to use the TCCON data format. In this section, we compare column averaged mole fractions of CO$_2$ and CH$_4$ of the CAMS models with TCCON retrievals. Data from the following TCCON sites has been used:
Izana (Blumenstock et al., 2017), Reunion (De Mazière et al., 2017), Bialystok (Deutscher et al., 2017), Manaus (Dubey et al., 2017), Four Corners (Dubey et al., 2017), Ascension (Feist et al., 2017), Anmeyondo (Goo et al., 2017), Darwin (Griffith et al., 2017), Wollongong (Griffith et al., 2017), Karlsruhe (Hase et al., 2017), Edwards (Iraci et al., 2017), Indianapolis (Iraci et al., 2017), Saga (Kawakami et al., 2017), Sodankyla (Kivi et al., 2017), Hefei (Liu et al., 2018), Tsukuba (Morino et al., 2017), Burgos (Morino et al., 2018), Rikubetsu (Morino et al., 2017), Bremen (Notholt et al., 2017), Spitsbergen (Notholt et al., 2017), Lauder (Sherlock et al., 2017, Pollard et al., 2019), Eureka (Strong et al., 2018), Garmisch (Sussmann et al., 2017), Zugspitze (Sussmann et al., 2018), Paris (Te et al., 2017), Orleans (Warneke et al., 2017), Park Falls (Wennberg et al., 2017), Caltech (Wennberg et al., 2017), Lamont (Wennberg et al., 2017), Jet Propulsion Laboratory (Wennberg et al., 2017), East Trout Lake (Wunch et al., 2017)

For the validation of the models in June, July and August the sites, for which data was available include Eureka, East Trout Lake, Orléans and Izana. Since TCCON PIs usually process the data in batches and the requirement within TCCON is to make the data publically available 1 year after the measurement, the availability of data for these reports is limited. Previously, data from Bialystok, Orléans and Réunion was always timely available for the validation of the CAMS models. The Bialystok site has stopped operation and the instrument has been transported to Cyprus, where measurements were started in September 2019. These data will become available for future reports. At Orléans a technical problem occurred in August, but the data stream is set up and fast data is delivered on a regular basis. At Réunion a technical problem is currently limiting the data availability. However, it can be expected that for future reports data from Cyprus, Orléans, Réunion and potentially a few other sites will be available.

Methane (CH$_4$)

Figure 9.2.1 shows the data for the last 4 years. The only data for the reporting period is from Eureka, East Trout Lake, Orléans and Izana. The data from these stations show that the model data continues to underestimate the CH$_4$ for these stations by 5-20 ppb as discussed in previous reports. Since for East Trout Lake only data for a few days is available its average differences are not comparable to the differences at the other sites (Fig 9.2.3).
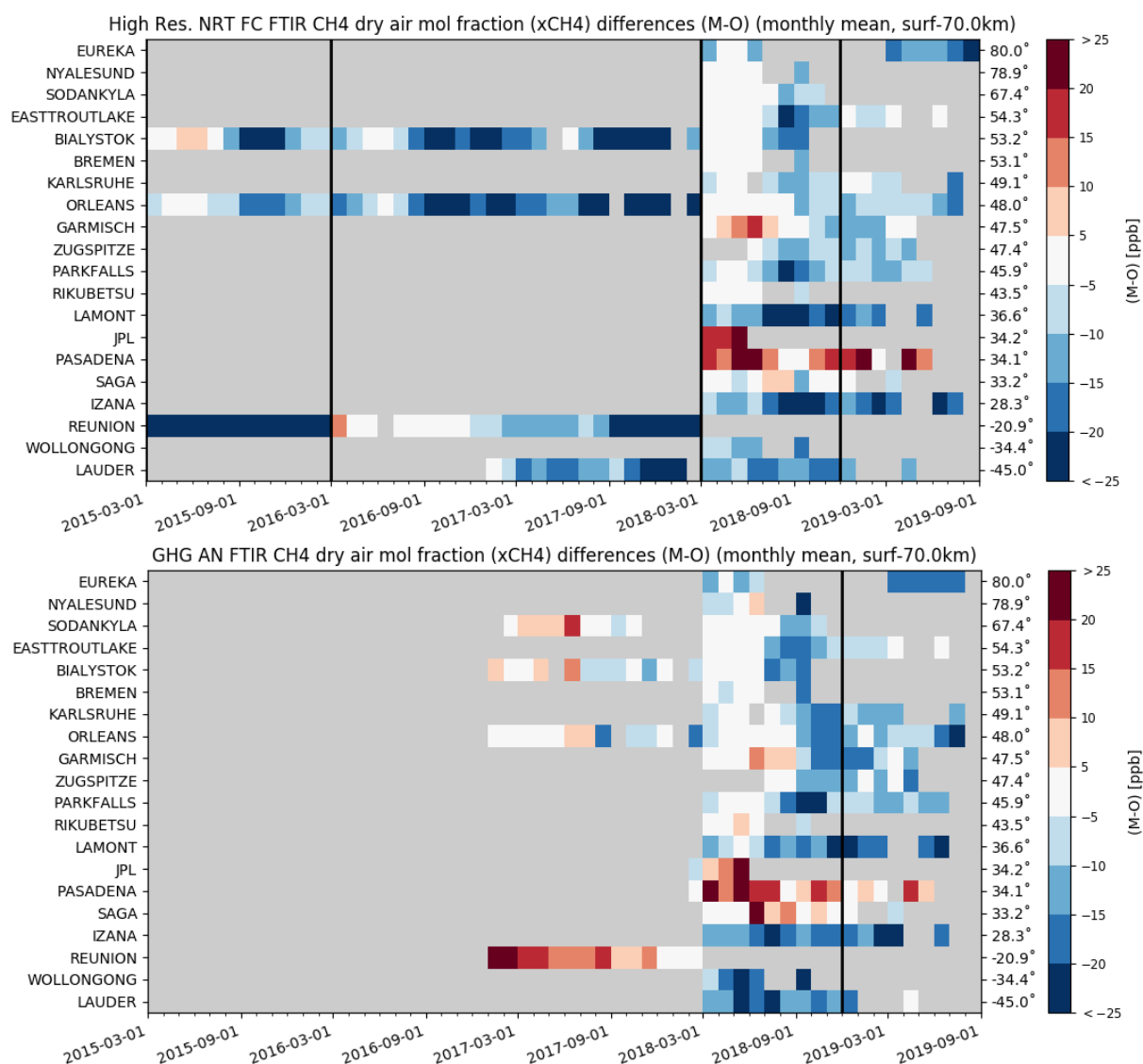
Figure 9.2.1: Monthly differences for the last 4 years (upper plot: high-resolution NRT, lower plot: GHG analysis). The stations are sorted by latitude (northern to southern hemisphere).
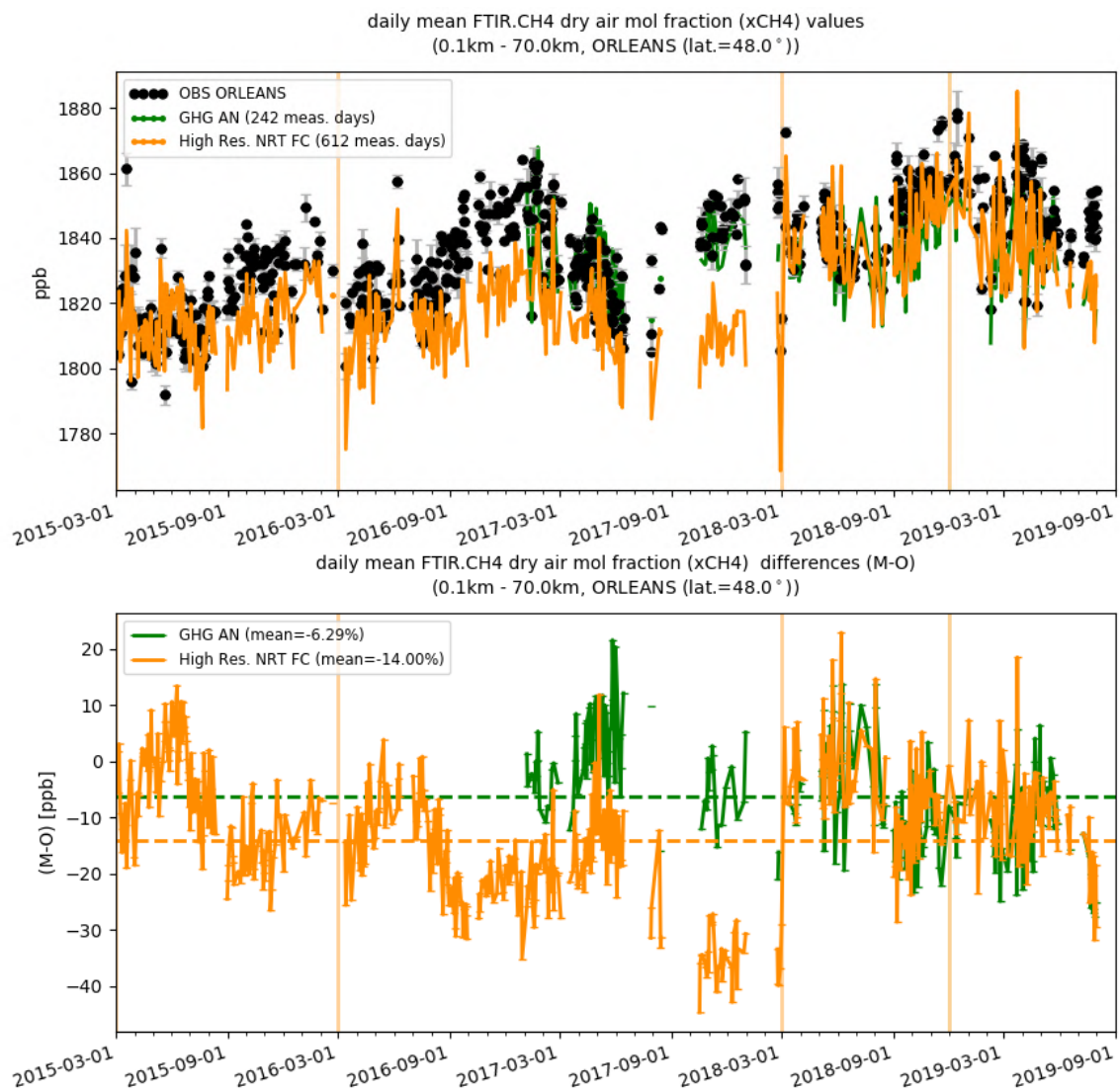
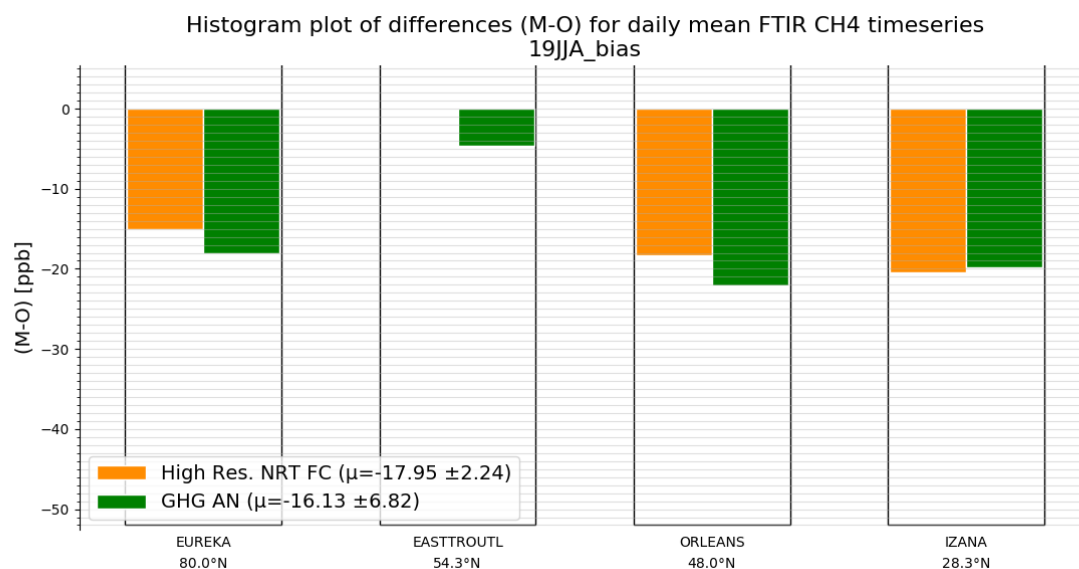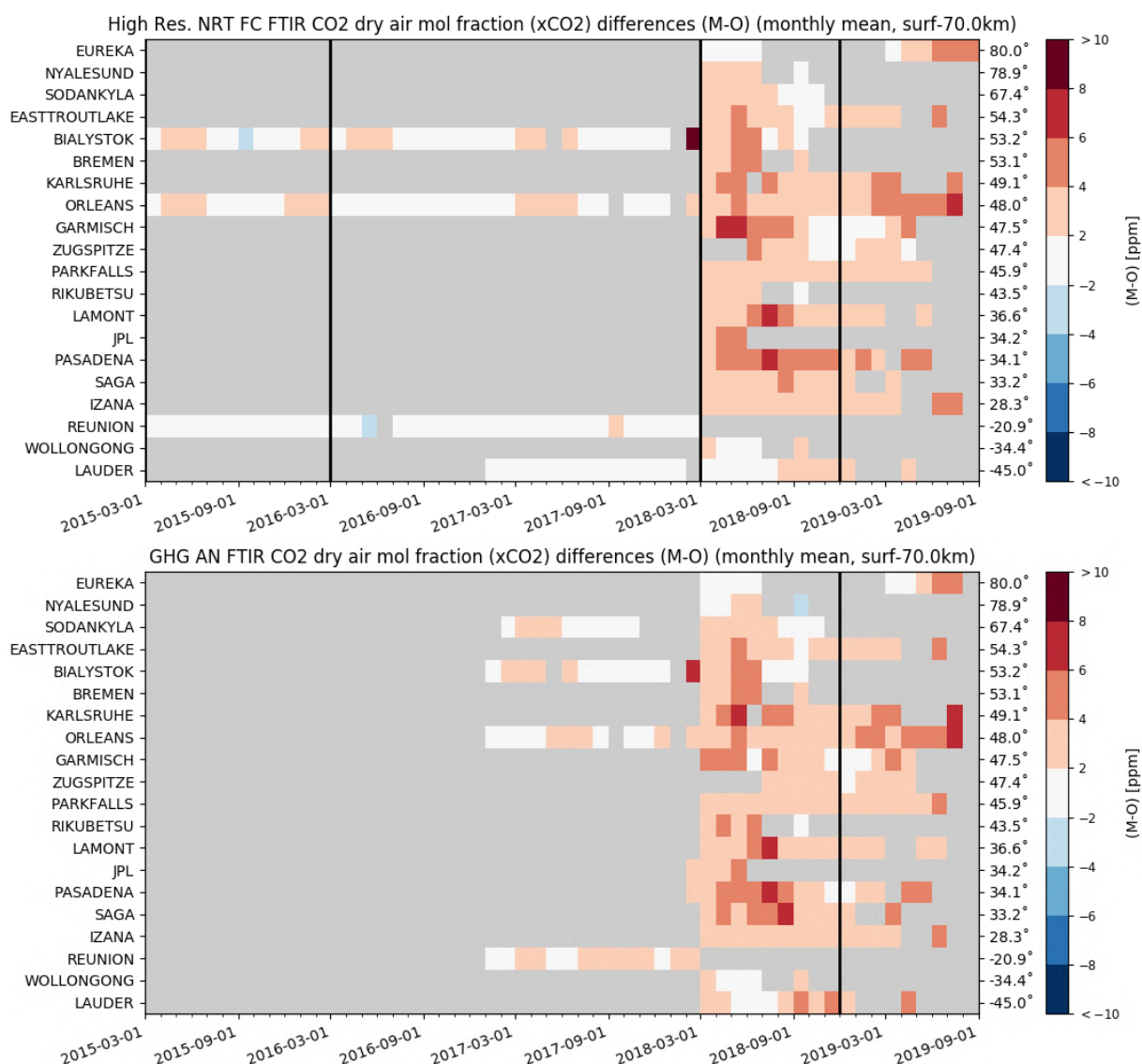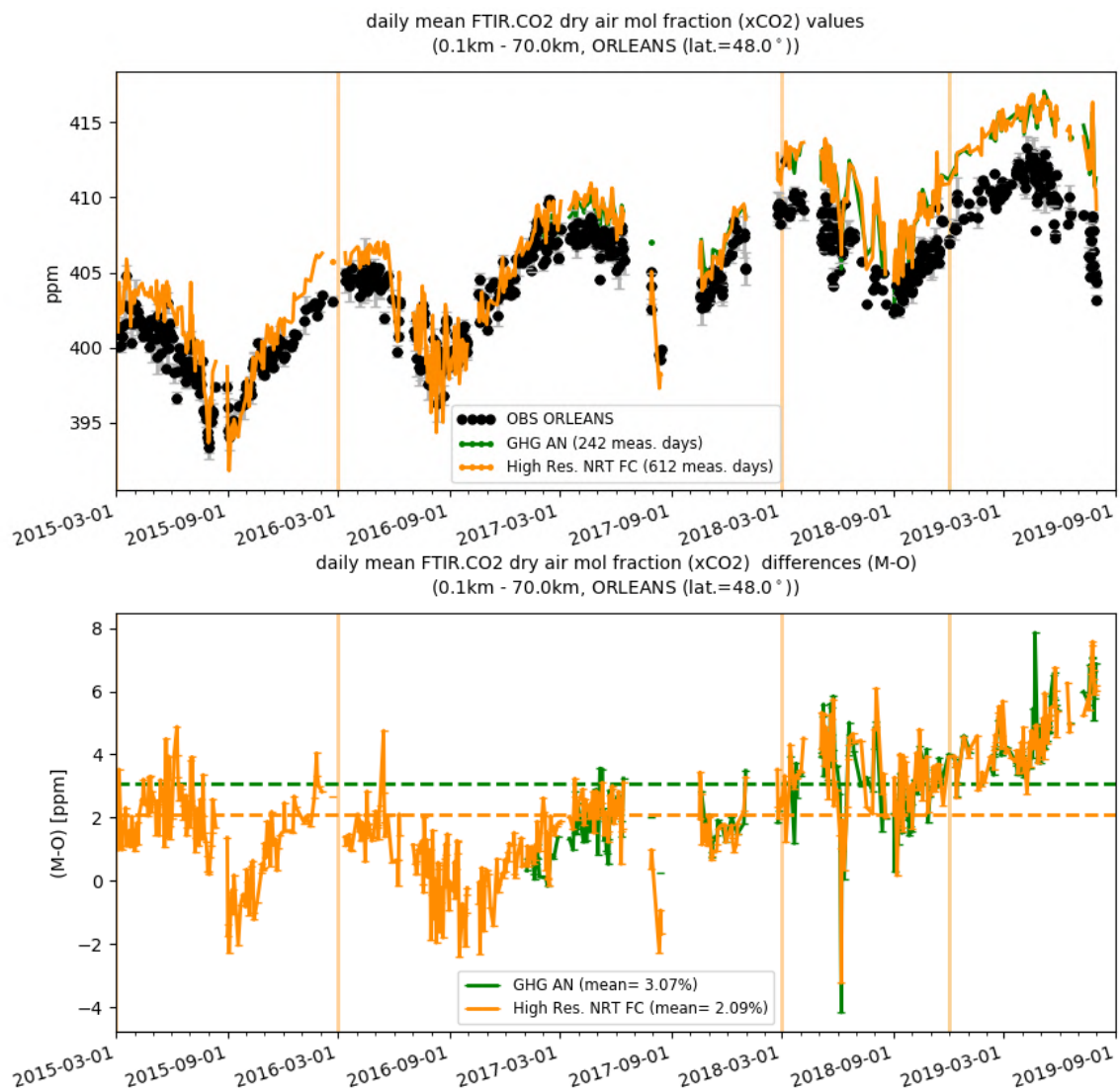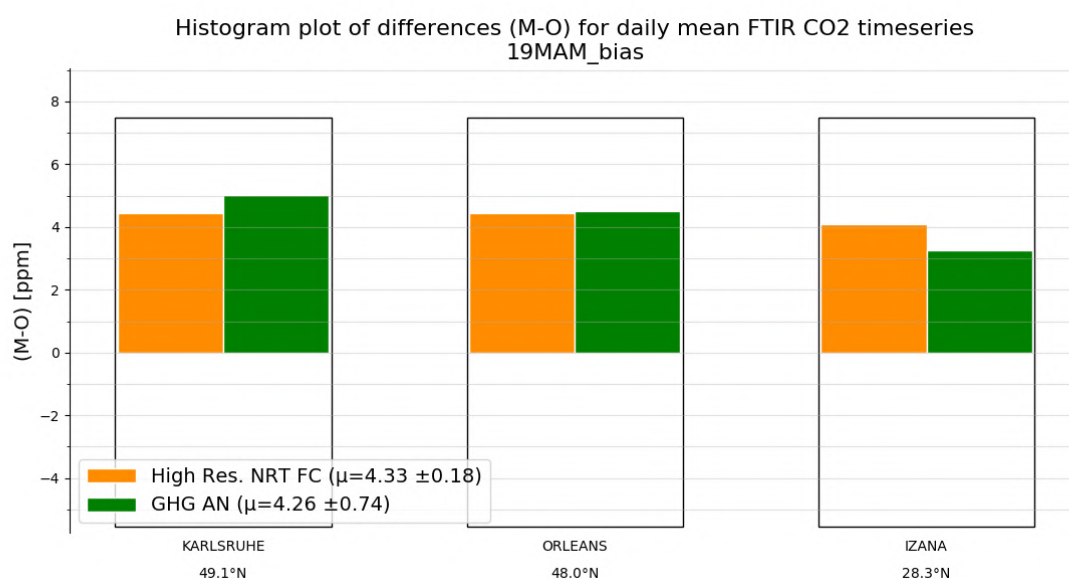Figure 9.2.2: Comparison of the XCH₄ model data with TCCON data at Orleans.

Figure 9.2.3: Differences during the reporting period.

Carbon dioxide (CO$_2$)

Figure 9.2.4 shows the data for the last 4 years. The only data available for the reporting period is from Eureka, East Trout Lake, Orléans and Izana. The data from these stations show that the model data continues to overestimate the CO$_2$. The comparison at Orléans shows that the overestimation is significantly higher than in previous years and reached up to 5-6 ppm (Fig. 9.2.5)

Figure 9.2.4: Monthly differences for the last 4 years (upper plot: high-resolution NRT; lower plot: GHG analysis). The stations are sorted by latitude (northern to southern hemisphere).

daily mean FTIR.CO2 dry air mol fraction (xCO2) values
(0.1km - 70.0km, ORLEANS (lat.=48.0°))

daily mean FTIR.CO2 dry air mol fraction (xCO2) differences (M-O)
(0.1km - 70.0km, ORLEANS (lat.=48.0°))

Figure 9.2.5: Comparison of the CO_2 model data with TCCON CO_2 at Orléans.

Figure 9.2.6: Differences during the reporting period

9.3 Validation against FTIR observations from the NDACC network

In this section, we compare the CH_4 profiles of the CAMS GHG products with FTIR measurements at different FTIR stations within the NDACC network. These ground-based, remote-sensing instruments are sensitive to the CH_4 abundance in the troposphere and lower stratosphere, i.e. between the surface and up to 25 km altitude. Tropospheric and stratospheric CH_4 columns are calculated from the FTIR profile data and used to validate corresponding columns obtained from the model data. A description of the instruments and applied methodologies can be found at http://nors.aeronomie.be. The typical uncertainty on the FTIR tropospheric column is 2%, while the uncertainty on the stratospheric column is 7.5%, adding together to a 3% uncertainty on the total column. The systematic uncertainty is large for the NDACC methane product mostly due to higher spectroscopic uncertainties.

Figure 9.3.1a/b (middle row) shows that the tropospheric columns of CH_4 agree well and only small differences appear between the analysis and the high-resolution model. In comparison with the measurement uncertainty, a slight underestimation is observed in the tropospheric columns which is in agreement with the TCCON results. The Paramaribo measurements have reduced sensitivity and the tropospheric/stratospheric split is not valid in this case.

The stratospheric columns (Figure 9.3.1a/b, bottom row) show a slight overestimation compared to the measurement uncertainty.

At some sites a seasonal change is observed in either the tropospheric or stratospheric concentrations. Due to the short time period, it is unclear if this is a recurring seasonal dependent model performance. In Figure 9.3.2 the tropospheric and stratospheric relative difference time series are plotted at Thule and St. Petersburg.
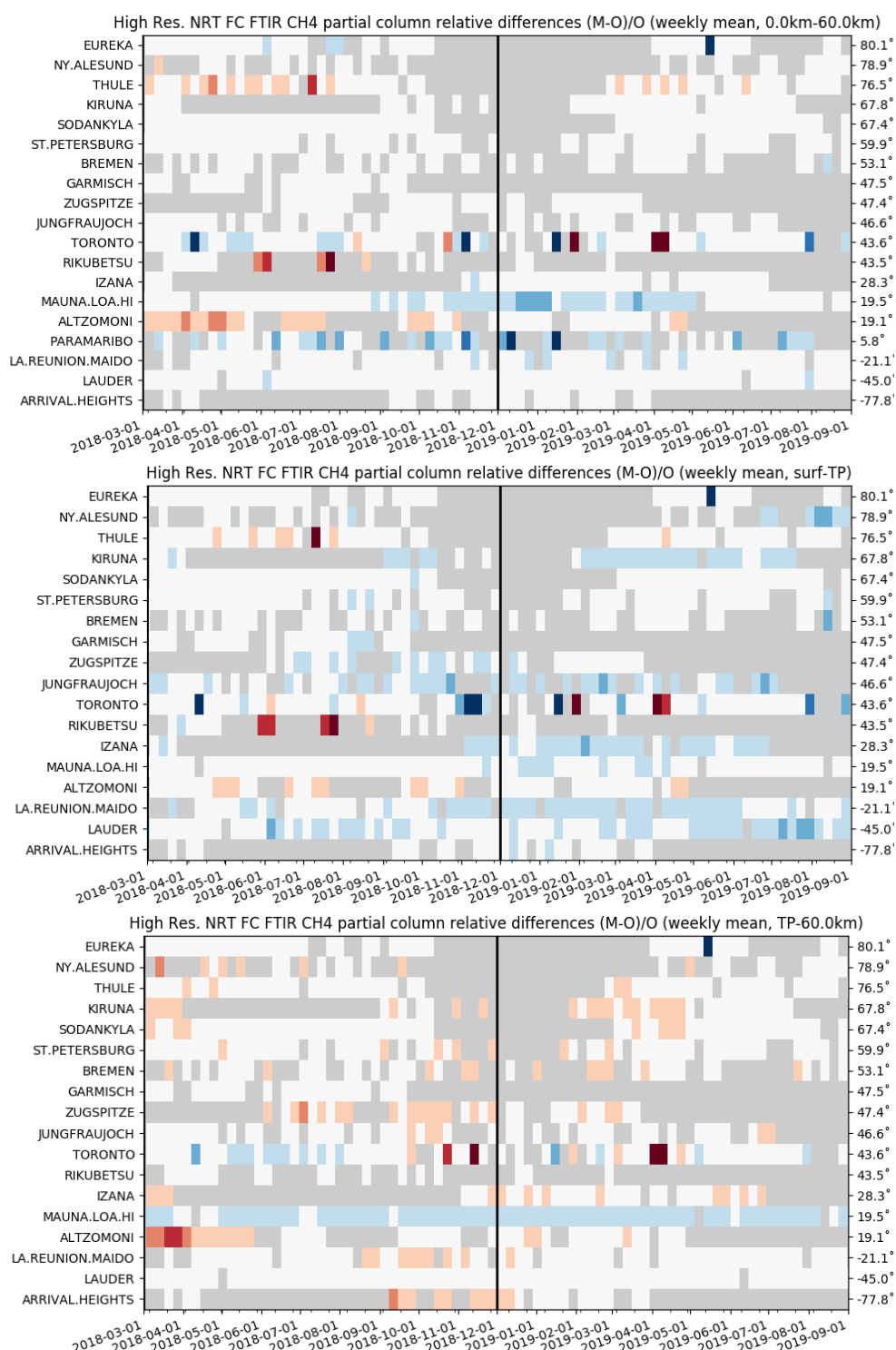
Figure 9.3.1a: Weekly mean relative bias for total (top row), tropospheric (middle row) and stratospheric CH₄ columns (bottom row) for the period March 2018 – August 2019 for high resolution forecasts. The overall uncertainty for the CH₄ total column measurements is approximately 4%. The overall uncertainty for the CH₄ total/tropospheric column measurements is approximately 2%, while the stratospheric uncertainty is 7.5% (color scale for the mosaic plots follows uncertainty scale).

GHG AN FTIR CH4 partial column relative differences (M-O)/O (weekly mean, 0.0km-60.0km)

GHG AN FTIR CH4 partial column relative differences (M-O)/O (weekly mean, surf-TP)

GHG AN FTIR CH4 partial column relative differences (M-O)/O (weekly mean, TP-60.0km)
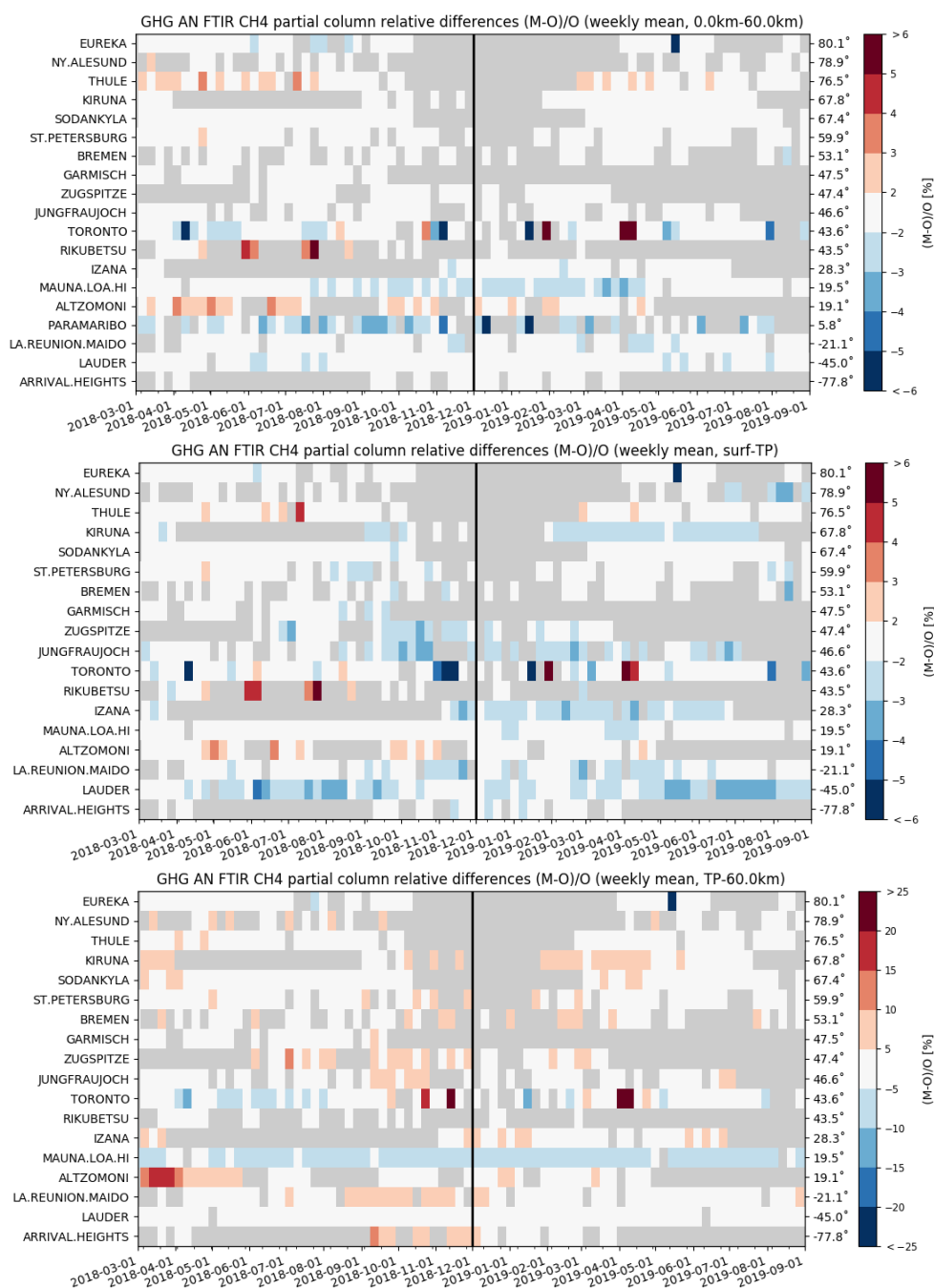
Figure 9.3.1b: Same as 9.3.1a, but for the analysis.

Figure 9.3.3 shows Taylor diagrams for the DJF time period and for a selected number of sites (many high latitude stations are not measuring during DJF): some stations have limited observations and should be treated with care. Assimilation has a small effect on the correlation coefficients for most sites: the average correlation is 0.65 for the analysis and 0.7 for the high-resolution forecast.

Figure 9.3.2: Daily mean of relative differences for tropospheric CH$_4$ columns (left) and stratospheric CH$_4$ columns (right) at Thule (top) and St Petersburg (bottom). At Thule the stratospheric column shows a reduced bias during the summer months, while at St Petersburg the tropospheric column performs worse during June-October.
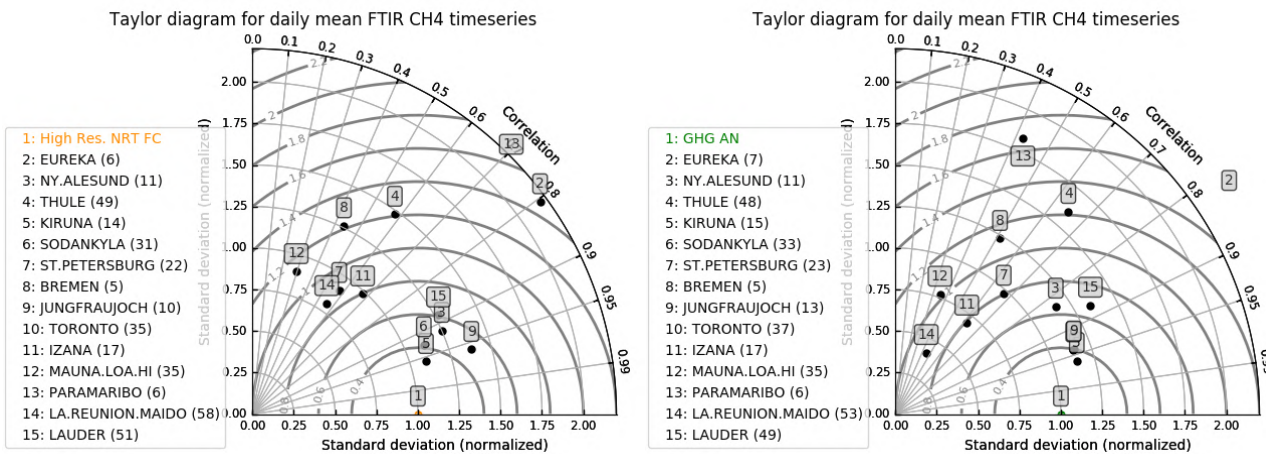
Figure 9.3.3: Taylor diagrams relating the standard deviations for the model /GB time series of total CH$_4$ column data and their correlation for the period 2019 JJA (the stations with a limited number of measurements should be ignored). All timeseries are normalized such that the std of the model column time series is 1. For the tropical sites Izana, Mauna Loa and Reunion (Maido) the assimilation decreases the ratio of the standard deviations of both time series significantly: the analysis methane columns are more variable compared to the high-resolution forecast.

10. Event studies

10.1 Dust hits Europe during June and July 2019

In summer 2019, the MODIS satellite instrument detected two dust outbreaks that reached Central Europe (see Figure 10.1.1). One outbreak occurred on 22-25 June, originating from Algeria and transported towards the Iberian Peninsula, hitting France and moving to the Eastern Mediterranean the next days. A second dust event was observed on 23-25 July which also originated in Algeria and Iberian Peninsula, Western France and arriving in the United Kingdom. Both events were nicely tracked by MODIS (see Figure 10.1.1, Figure 10.1.2 and Figure 10.1.3) and the AERONET sun photometers in the central Iberian Peninsula (see Figure 7.4.3).
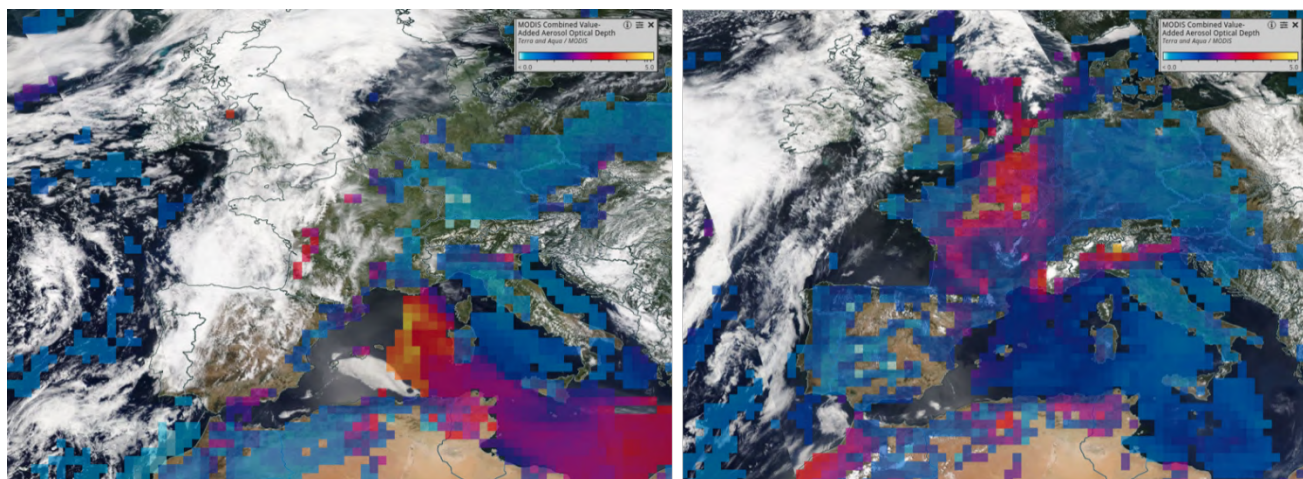
Figure 10.1.1: Daily AOD at 550nm composite of NASA MODIS Terra/Aqua on 24th June (left image) and 24th July (right image) 2019 over Europe. These images are a zoom of the images included in the comparison with CAMS o-suite in Figure 10.1.2 and Figure 10.1.3.

CAMS AOD o-suite did timely reproduce the spatial distribution of the two dust plumes as shown in the comparison with MODIS/Aqua AOD comparison (see Figure 10.1.2 and Figure 10.1.3) despite the model tendency to overestimate the observed maximum values, in particular the second event in which the Madrid AERONET site in Central Spain (Figure 10.1.4) shows values of AOD up to 0.4 and the o-suite predicts AOD values around 0.6 on 24th July. This second dust outbreak was associated with a deep depression. During the second event, the o-suite is predicting PM10 values over 100μg/m^3 in large parts of the Iberian Peninsula and France (Figure 10.1.3) and achieving values up to 50μg/m^3 in United Kingdom (Figure 10.1.5).
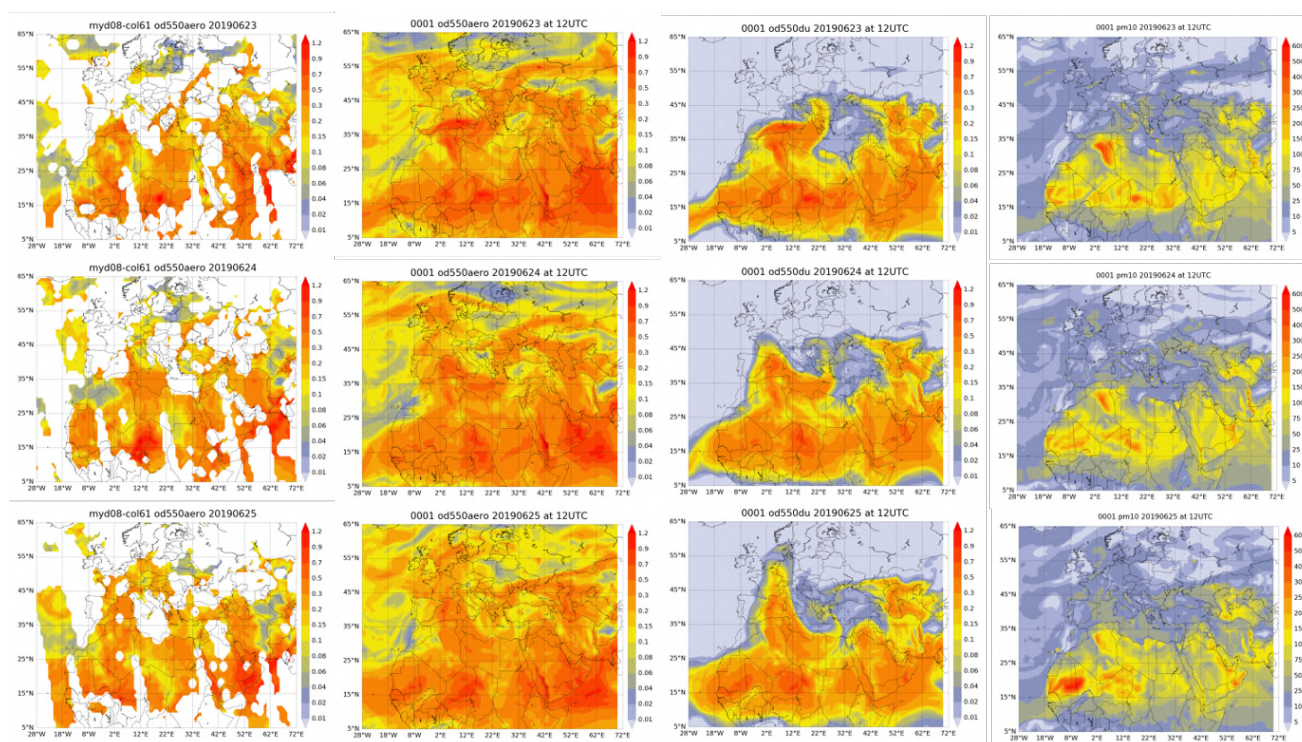
Figure 10.1.2. AOD from MODIS/Aqua Collection 6.1 Daily Level 3 product as well as AOD (second column), DOD (third column) and PM10 (last column) at 12UTC from the CAMS o-suite for 23-25 June 2019.
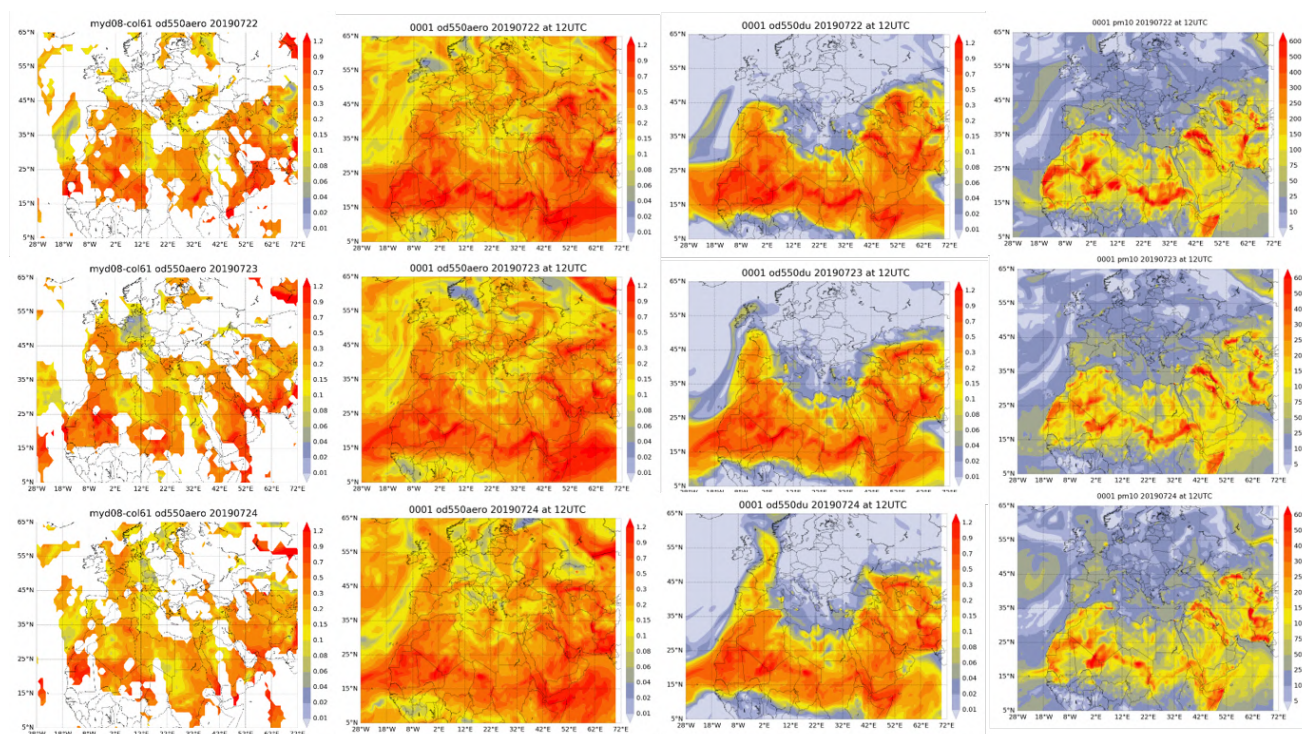
Figure 10.1.3. AOD from MODIS/Aqua Collection 6.1 Daily Level 3 product as well as AOD (second column), DOD (third column) and PM10 (last column) at 12UTC from the CAMS o-suite for 22-24 July 2019.
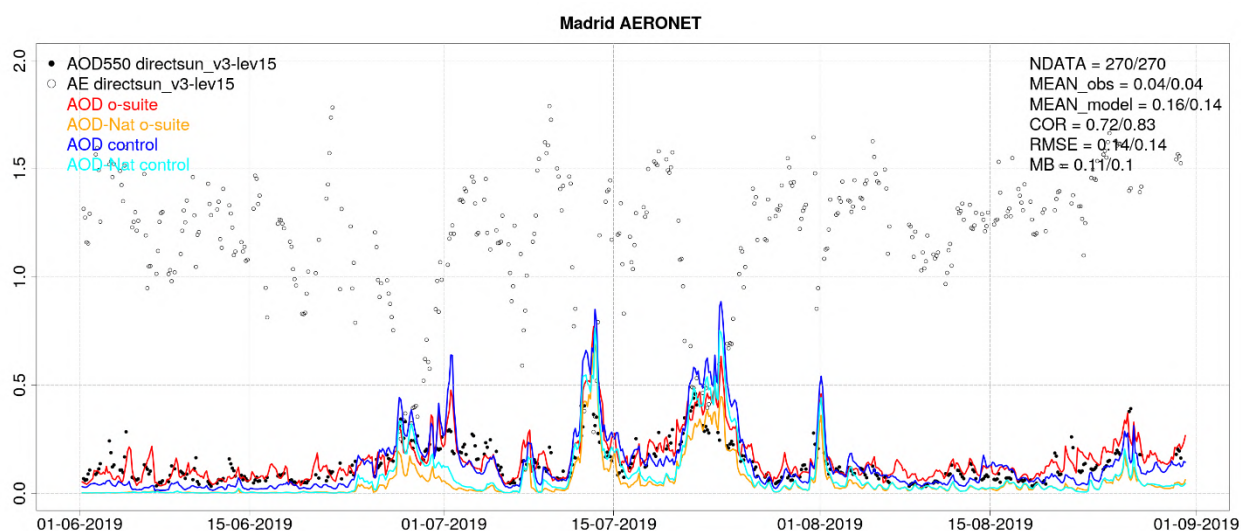
Figure 10.1.4: AOD from AERONET (black dot), AOD o-suite (red line), AOD control (blue line), AOD-Nat o-suite (orange line), AOD-Nat control (cyan line), for the study period Madrid (Central Iberian Peninsula). AOD-Nat corresponds to the natural aerosol optical depth that includes dust and sea-salt. Skill scores per each individual site and model (o-suite/control) are shown in the upper right corner (including NDATA: available 3-hourly values used for the calculations, mean of observations, mean of the model, correlation (COR), RMSE, MB).
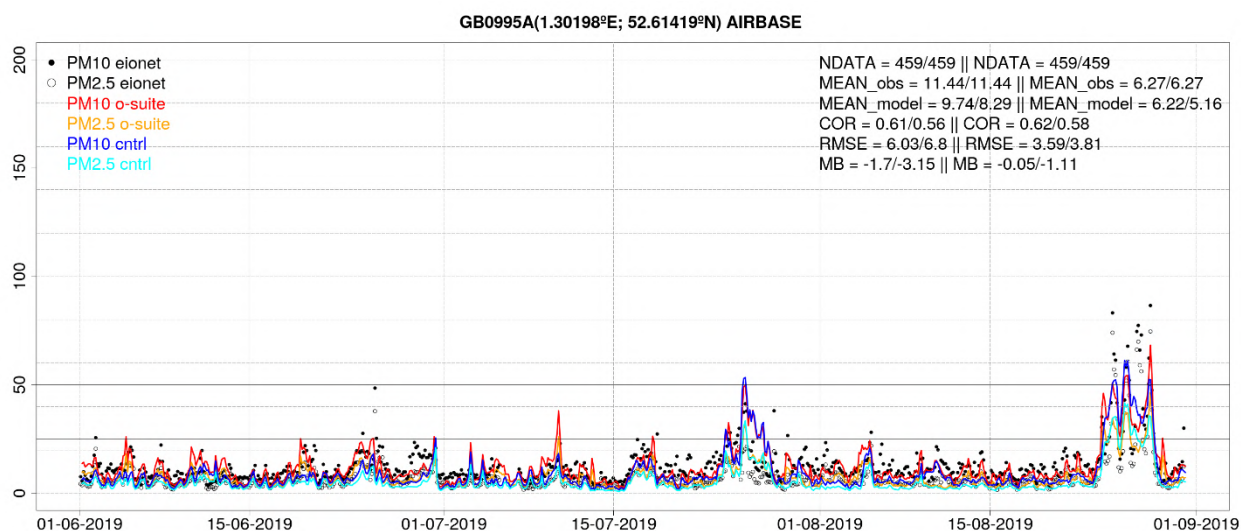
Figure 10.1.5: PM10 and PM2.5 Airbase observations (black and grey dots, respectively), PM10 and PM2.5 o-suite (red and orange lines, respectively) and PM10 and PM2.5 control (blue and cyan lines, respectively) for the study period over GB0995A (United Kingdom).

10.2 Syberian fire event and plumes over the Arctic

A large fire event centred at approximately 110°E and 60°N in Siberia north of the Baikal Lake started on 6 August with large CO emissions for about one week. There was a northwards transport of CO, which reached the Arctic Ocean on 9 August. A plume travelled across the Arctic Ocean and reached north Greenland on 13 August. Elevated concentrations were also found over northwest Greenland the Canadian Arctic the following days. The plume was dispersed down the East Greenland coast and spiralled around Svalbard, re-entering the Arctic Ocean east of Svalbard the following days (Figure 10.2.1). The episode is generally well captured by the o-suite, although the total column CO is underestimated (Figure 10.2.2). The control data is even further underestimating the CO levels, but the transport episode can still be seen in the control data. The episode is not captured by any surface measurements in Greenland and Svalbard. A peak in total column CO measured with NDACC FTIR at Thule on 16-17 August may be associated with this transport event. Elevated concentrations are also predicted by the o-suite, although the model underestimates the levels. (Figure 10.2.3).
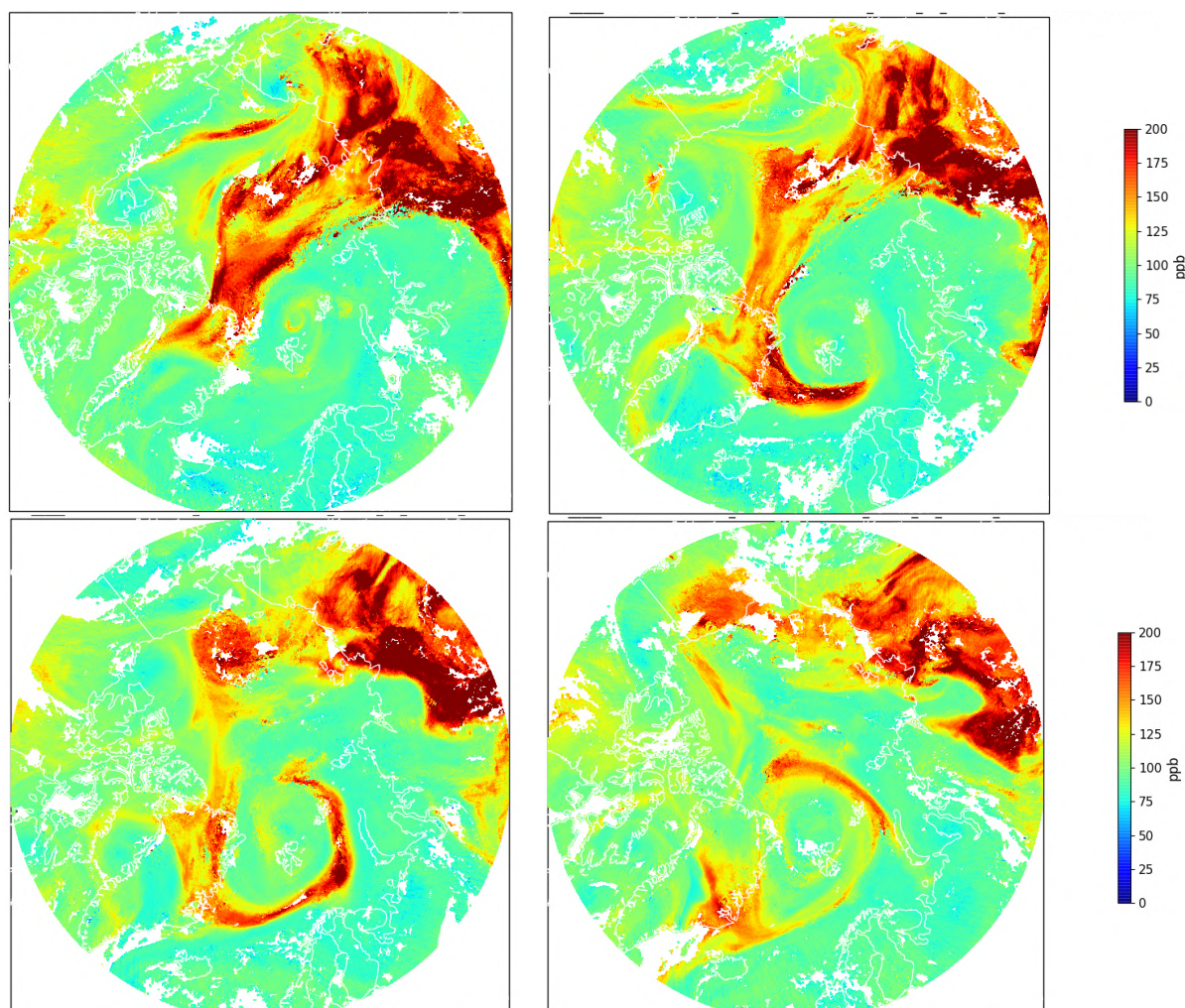
Figure 10.2.1: Total column CO as measured from Sentinel-5P on 14 (top-left), 15 (top-right), 16 (bottom-left) and 17 August (bottom-right).
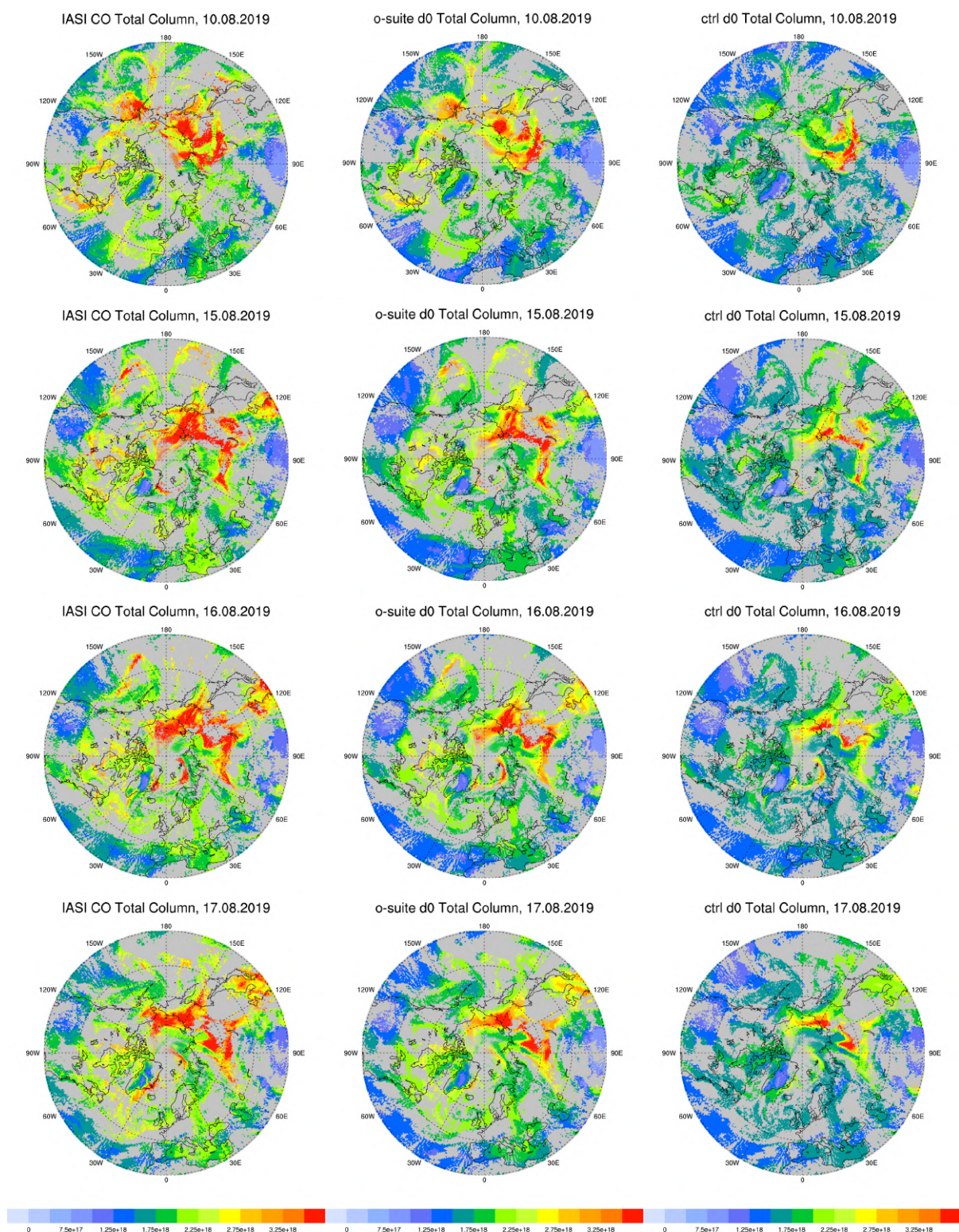
Figure 10.2.2: Total column CO as measured by IASI (left) and predicted by the o-suite (middle) and control run (right) for 10 August (top row), 15 August (second row), 16 August (third row), and 17 August (bottom).
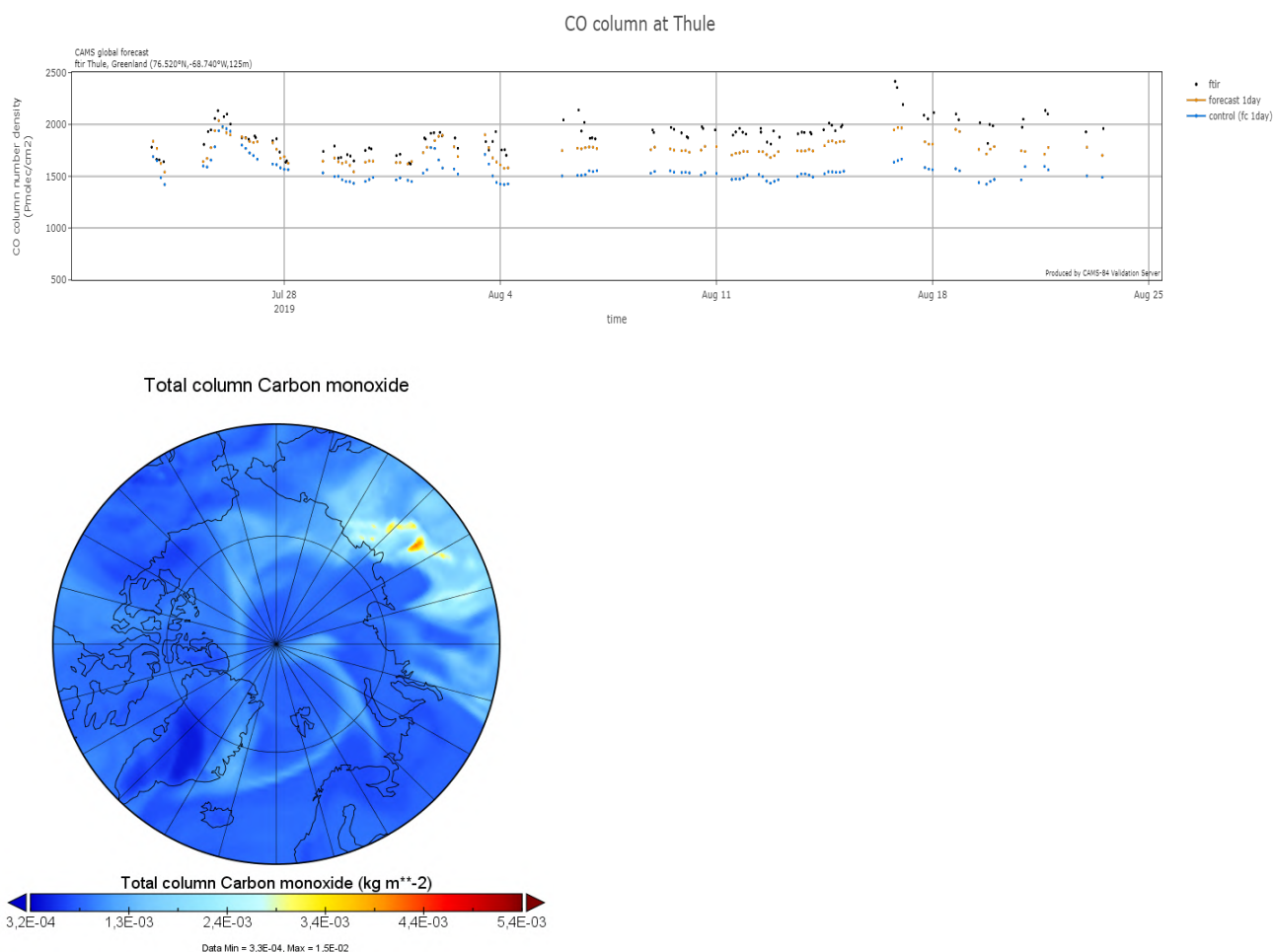
Figure 10.2.3: Total column CO measured by the NDACC FTIR at Thule, Greenland between 23 July and 25 August, with a distinct peak on 16 August (top) and predicted total column CO by the o-suite (bottom) for 16 August at 21:00.

11. References

Agusti-Panareda, A., Monitoring upgrades of analysis/forecast system, MACC-III Deliverable D44.04, June 2015.

Basart, S, A. Benedictow, Y. Bennouna, A.-M. Blechschmidt, S. Chabrillat, Y. Christophe, E. Cuevas, H. J. Eskes, K. M. Hansen, O. Jorba, J. Kapsomenakis, B. Langerock, T. Pay, A. Richter, N. Sudarchikova, M. Schulz, A. Wagner, C. Zerefos, Upgrade verification note for the CAMS near-real time global atmospheric composition service: Evaluation of the e-suite for the CAMS upgrade of July 2019, Copernicus Atmosphere Monitoring Service (CAMS) report, CAMS84_2018SC1_D3.2.1-201907_esuite_v1.pdf, July 2019.

Bergamaschi, P., Frankenberg, C., Meirink, J. F., Krol, M., Villani, M. G., Houweling, S., Dentener, F., Dlugokencky, E. J., Miller, J. B., Gatti, L. V., Engel, A., and Levin, I.: Inverse modeling of global and regional CH4 emissions using SCIAMACHY satellite retrievals, J. Geophys. Res., 114, D22301, doi:10.1029/2009JD012287, 2009.

Benedetti, A., J.-J. Morcrette, O. Boucher, A. Dethof, R. J. Engelen, M. Fisher, H. Flentjes, N. Huneeus, L. Jones, J. W. Kaiser, S. Kinne, A. Mangold, M. Razinger, A. J. Simmons, M. Suttie, and the GEMS-AER team: Aerosol analysis and forecast in the ECMWF Integrated Forecast System. Part II : Data assimilation, J. Geophys. Res., 114, D13205, doi:10.1029/2008JD011115, 2009.

Boussetta, S., Balsamo, G., Beljaars, A., Agusti-Panareda, A., Calvet, J.-C., Jacobs, C., van den Hurk, B., Viterbo, P., Lafont, S., Dutra, E., Jarlan, L., Balzarolo, M., Papale, D., and van der Werf, G.: Natural carbon dioxide exchanges in the ECMWF Integrated Forecasting System: implementation and offline validation, J. Geophys. Res.-Atmos., 118, 1–24, doi: 10.1002/jgrd.50488, 2013.

Braathen, WMO Arctic Ozone Bulletin No 1/2016, DOI:10.13140/RG.2.1.4929.6403, 2016.

Cammas, J.P., Brioude J., Chaboureau J.-P., Duron J., Mari C., Mascart P., Nédélec P., Smit H., Pätz H.-W., Volz-Thomas A., Stohl A., and Fromm M., Injection in the lower stratosphere of biomass fire emissions followed by long-range transport: a MOZAIC case study. Atmos. Chem. Phys., 9, 5829-5846, 2009

Cariolle, D. and Teyssèdre, H.: A revised linear ozone photochemistry parameterization for use in transport and general circulation models: multi-annual simulations, Atmos. Chem. Phys., 7, 2183-2196, doi:10.5194/acp-7-2183-2007, 2007.

Dee, D. P. and S. Uppala, Variational bias correction of satellite radiance data in the ERA-Interim reanalysis. Quart. J. Roy. Meteor. Soc.,135, 1830-1841, 2009.

Deeter, M. N., Emmons, L. K., Edwards, D. P., Gille, J. C., and Drummond, J. R.: Vertical resolution and information content of CO profiles retrieved by MOPITT, Geophys. Res. Lett., 31, L15112, doi:10.1029/2004GL020235, 2004.

Deeter, M. N., et al. (2010), The MOPITT version 4 CO product: Algorithm enhancements, validation, and long-term stability, J. Geophys. Res., 115, D07306, doi:10.1029/2009JD013005.

Dentener, F., et al., 2006: Emissions of primary aerosol and precursor gases in the years 2000 and 1750 prescribed data-sets for AeroCom, Atmos. Chem. Phys., 6, 4321 – 4344.

Deshler, T., J.L. Mercer, H.G.J. Smit, R. Stubi, G. Levrat, B.J. Johnson, S.J. Oltmans, R. Kivi, A.M. Thompson, J. Witte, J. Davies, F.J. Schmidlin, G. Brothers, T. Sasaki (2008) Atmospheric comparison of electrochemical cell ozonesondes from different maufacturers, and with different cathode solution strengths: The Balloon Experiment on Standards for Ozonsondes. J. Geophys. Res.113, D04307, doi:10.1029/2007JD008975

Dupuy, E., et al.: Validation of ozone measurements from the Atmospheric Chemistry Experiment (ACE), Atmos. Chem. Phys., 9, 287-343, doi:10.5194/acp-9-287-2009, 2009.

Elbern, H., Schwinger, J., Botchorishvili, R.: Chemical state estimation for the middle atmosphere by four-dimensional variational data assimilation: System configuration. Journal of Geophysical Research (Atmospheres) 115, 6302, 2010.

Emmons, L. K., D. P. Edwards, M. N. Deeter, J. C. Gille, T. Campos, P. Nédélec, P. Novelli, and G. Sachse, Measurements of Pollution In The Troposphere (MOPITT) validation through 2006 Atmos. Chem. Phys., 9, 1795-1803, 2009

Errera, Q., Daerden, F., Chabrillat, S., Lambert, J. C., Lahoz, W. A., Viscardy, S., Bonjean, S., and Fonteyn, D., 4D-Var Assimilation of MIPAS chemical observations: ozone and nitrogen dioxide analyses, Atmos. Chem. Phys., 8, 6169-6187, 2008.

Errera, Q. and Ménard, R.: Technical Note: Spectral representation of spatial correlations in variational assimilation with grid point models and application to the belgian assimilation system for chemical observations (BASCOE), Atmos. Chem. Phys. Discuss., 12, 16763-16809, doi:10.5194/acpd-12-16763-2012, 2012.

Eskes, H.J., S. Basart, A. Benedictow, Y. Bennouna, A.-M. Blechschmidt, S. Chabrillat, Y. Christophe, E. Cuevas, J. Douros, H. Flentje, K. M. Hansen, J. Kapsomenakis, B. Langerock, M. Ramonet, A. Richter, M. Schulz, N. Sudarchikova, A. Wagner, T. Warneke, C. Zerefos, Observations characterisation and validation methods document, Copernicus Atmosphere Monitoring Service (CAMS) report, CAMS84_2015SC3_D.84.8.1.1-2018_observations_v3.pdf, October 2018 (2018a). Available from: http://atmosphere.copernicus.eu/user-support/validation/verification-global-services

Eskes, H. J., S. Basart, A. Benedictow, Y. Bennouna, A.-M. Blechschmidt, S. Chabrillat, Y. Christophe, H. Clark, E. Cuevas, K. M. Hansen, U. Im, J. Kapsomenakis, B. Langerock, K. Petersen, M. Schulz, A. Wagner, C. Zerefos, Upgrade verification note for the CAMS near-real time global atmospheric composition service, Copernicus Atmosphere Monitoring Service (CAMS) report, CAMS84_2015SC3_D84.3.1.5_201802_esuite_v1.pdf, February 2018 (2018b)

Eskes et al., Upgrade verification note for the CAMS near-real time global atmospheric composition service, Addendum July 2018, CAMS84_2015SC3_D84.3.1.5_201802_esuite_v1.pdf (2018c).

Flemming, J., Huijnen, V., Arteta, J., Bechtold, P., Beljaars, A., Blechschmidt, A.-M., Diamantakis, M., Engelen, R. J., Gaudel, A., Inness, A., Jones, L., Josse, B., Katragkou, E., Marecal, V., Peuch, V.-H., Richter, A., Schultz, M. G., Stein, O., and Tsikerdekis, A.: Tropospheric chemistry in the Integrated Forecasting System of ECMWF, Geosci. Model Dev., 8, 975-1003, doi:10.5194/gmd-8-975-2015, 2015.

Flemming, J., Benedetti, A., Inness, A., Engelen, R. J., Jones, L., Huijnen, V., Remy, S., Parrington, M., Suttie, M., Bozzo, A., Peuch, V.-H., Akritidis, D., and Katragkou, E.: The CAMS interim Reanalysis of Carbon Monoxide, Ozone and Aerosol for 2003–2015, Atmos. Chem. Phys., 17, 1945-1983, doi:10.5194/acp-17-1945-2017, 2017.

Franco, B., et al., Retrievals of formaldehyde from ground-based FTIR and MAX-DOAS observations at the Jungfraujoch station and comparisons with GEOS-Chem and IMAGES model simulations, Atmos. Meas. Tech., 8, 1733-1756, 2015

Gielen, C., Van Roozendael, M., Hendrick, F., Pinardi, G., Vlemmix, T., De Bock, V., De Backer, H., Fayt, C., Hermans, C., Gillotay, D., and Wang, P.: A simple and versatile cloud-screening method for MAX-DOAS retrievals, Atmos. Meas. Tech., 7, 3509-3527, doi:10.5194/amt-7-3509-2014, 2014.

Granier, C. et al.: Evolution of anthropogenic and biomass burning emissions of air pollutants at global and regional scales during the 1980–2010 period. Climatic Change (109), 2011

Holben, B. N., Eck, T. F., Slutsker, I., Tanré, D., Buis, J. P., Setzer, A., Vermote, E., Reagan, J. A., Kaufman, Y. J., Nakajima, T., Lavenu, F., Jankowiak, I., and Smirnov A.: AERONET – a federated instrument network and data archive for aerosol characterization, Remote Sens. Environ., 66, 1–16, 5529, 5533, 5537, 5544, 1998.

Hommel, R., Eichmann, K.-U., Aschmann, J., Bramstedt, K., Weber, M., von Savigny, C., Richter, A., Rozanov, A., Wittrock, F., Khosrawi, F., Bauer, R., and Burrows, J. P.: Chemical ozone loss and ozone mini-hole event during the Arctic winter 2010/2011 as observed by SCIAMACHY and GOME-2, Atmos. Chem. Phys., 14, 3247-3276, doi:10.5194/acp-14-3247-2014, 2014.

Huijnen, V., et al.: The global chemistry transport model TM5: description and evaluation of the tropospheric chemistry version 3.0, Geosci. Model Dev., 3, 445-473, doi:10.5194/gmd-3-445-2010, 2010.

Inness, A., Blechschmidt, A.-M., Bouarar, I., Chabrillat, S., Crepulja, M., Engelen, R. J., Eskes, H., Flemming, J., Gaudel, A., Hendrick, F., Huijnen, V., Jones, L., Kapsomenakis, J., Katragkou, E., Keppens, A., Langerock, B., de Mazière, M., Melas, D., Parrington, M., Peuch, V. H., Razinger, M., Richter, A., Schultz, M. G., Suttie, M., Thouret, V., Vrekoussis, M., Wagner, A., and Zerefos, C.: Data assimilation of satellite-retrieved ozone, carbon monoxide and nitrogen dioxide with ECMWF's Composition-IFS, Atmos. Chem. Phys., 15, 5275-5303, doi:10.5194/acp-15-5275-2015, 2015.

Janssens-Maenhout, G., Dentener, F., Aardenne, J. V., Monni, S., Pagliari, V., Orlandini, L., Klimont, Z., Kurokawa, J., Akimoto, H., Ohara, T., Wankmueller, R., Battye, B., Grano, D., Zuber, A., and Keating, T.: EDGAR-HTAP: a Harmonized Gridded Air Pollution Emission Dataset Based on National Inventories, JRC68434, EUR report No EUR 25 299–2012, ISBN 978-92-79- 23122-0, ISSN 1831-9424, European Commission Publications Office, Ispra (Italy), 2012.

Jaross, G., Bhartia, P.K., Chen, G., Kowitt, M., Haken, M., Chen, Z., Xu, Ph., Warner, J., Kelly, T. : OMPS Limb Profiler instrument performance assessment, J. Geophys. Res. Atmos 119, 2169-8996, 2014.

Kaiser, J. W., Heil, A., Andreae, M. O., Benedetti, A., Chubarova, N., Jones, L., Morcrette, J.-J., Razinger, M., Schultz, M. G., Suttie, M., and van der Werf, G. R.: Biomass burning emissions estimated with a global fire assimilation system based on observed fire radiative power, Biogeosciences, 9, 527-554, doi:10.5194/bg-9-527-2012, 2012.

Kramarova, N. A., Nash, E. R., Newman, P. A., Bhartia, P. K., McPeters, R. D., Rault, D. F., Seftor, C. J., Xu, P. Q., and Labow, G. J.: Measuring the Antarctic ozone hole with the new Ozone Mapping and Profiler Suite (OMPS), Atmos. Chem. Phys., 14, 2353-2361, doi:10.5194/acp-14-2353-2014, 2014.

Lahoz, W. A., Errera, Q., Viscardy, S., and Manney G. L., The 2009 stratospheric major warming described from synergistic use of BASCOE water vapour analyses and MLS observations, Atmos. Chem. Phys. 11, 4689-4703, 2011

Lambert, A, et al., Aura Microwave Limb Sounder Version 3.4 Level-2 near real-time data user guide, http://disc.sci.gsfc.nasa.gov/Aura/data-holdings/MLS/documents/NRT-user-guide-v34.pdf

Langerock, B., De Mazière, M., Hendrick, F., Vigouroux, C., Desmet, F., Dils, B., and Niemeijer, S.: Description of algorithms for co-locating and comparing gridded model data with remote-sensing observations, Geosci. Model Dev., 8, 911-921, doi:10.5194/gmd-8-911-2015, 2015.

Lefever, K., van der A, R., Baier, F., Christophe, Y., Errera, Q., Eskes, H., Flemming, J., Inness, A., Jones, L., Lambert, J.-C., Langerock, B., Schultz, M. G., Stein, O., Wagner, A., and Chabrillat, S.: Copernicus stratospheric ozone service, 2009–2012: validation, system intercomparison and roles of input data sets, Atmos. Chem. Phys., 15, 2269-2293, doi:10.5194/acp-15-2269-2015, 2015.

Liu, Z., et al., Exploring the missing source of glyoxal (CHOCHO) over China, Geophys. Res. Lett., 39, L10812, doi: 10.1029/2012GL051645, 2012

Massart, S., Flemming, J., Cariolle, D., Jones, L., High resolution CO tracer forecasts, MACC-III Deliverable D22.04, May 2015, available from http://www.gmes-atmosphere.eu/documents/macciii/deliverables/grg

Morcrette, J.-J., O. Boucher, L. Jones, D. Salmond, P. Bechtold, A. Beljaars, A. Benedetti, A. Bonet, J. W. Kaiser, M. Razinger, M. Schulz, S. Serrar, A. J. Simmons, M. Sofiev, M. Suttie, A. M. Tompkins, and A. Untch: Aerosol analysis and forecast in the ECMWF Integrated Forecast System. Part I: Forward modelling, J. Geophys. Res., 114, D06206, doi:10.1029/2008JD011235, 2009.

Rémy, S., Kipling, Z., Flemming, J., Boucher, O., Nabat, P., Michou, M., Bozzo, A., Ades, M., Huijnen, V., Benedetti, A., Engelen, R., Peuch, V.-H., and Morcrette, J.-J.: Description and evaluation of the tropospheric aerosol scheme in the European Centre for Medium-Range Weather Forecasts (ECMWF) Integrated Forecasting System (IFS-AER, cycle 45R1), Geosci. Model Dev., 12, 4627–4659, https://doi.org/10.5194/gmd-12-4627-2019, 2019.

Richter, A., Burrows, J. P., Nüß, H., Granier, C, Niemeier, U.: Increase in tropospheric nitrogen dioxide over China observed from space, Nature, 437, 129-132, doi: 10.1038/nature04092, 2005

Richter, A., Begoin, M., Hilboll, A., and Burrows, J. P.: An improved NO2 retrieval for the GOME-2 satellite instrument, Atmos. Meas. Tech. , **4**, 1147-1159, doi:10.5194/amt-4-1147-2011, 2011

Schulz, M., Y. Christophe, M. Ramonet, Wagner, A., H.J. Eskes, S. Basart, A. Benedictow, Y. Bennouna, A.-M. Blechschmidt, S. Chabrillat, E. Cuevas, A. El- Yazidi, H. Flentje, K.M. Hansen, U. Im, J. Kapsomenakis, B. Langerock, A. Richter, N. Sudarchikova, V. Thouret, T. Warneke, C. Zerefos, Validation report of the CAMS near-real-time global atmospheric composition service: Period December 2018 -February 2019, Copernicus Atmosphere Monitoring Service (CAMS) report, CAMS84_2018SC1_D1.1.1_DJF2019_v1.pdf, June 2019, doi:10.24380/7th6-tk72.

Sindelarova, K., Granier, C., Bouarar, I., Guenther, A., Tilmes, S., Stavrakou, T., Müller, J.-F., Kuhn, U., Stefani, P., and Knorr, W.: Global data set of biogenic VOC emissions calculated by the MEGAN model over the last 30 years, Atmos. Chem. Phys., 14, 9317-9341, doi:10.5194/acp-14-9317-2014, 2014.

Smit, H.G.J., W. Straeter, B.J. Johnson, S.J. Oltmans, J. Davies, D.W. Tarasick, B. Hoegger, R. Stubi, F.J. Schmidlin, T. Northam, A.M. Thompson, J.C. Witte, I. Boyd: Assessment of the performance of ECC-ozonesondes under quasi-flight conditions in the environmental simulation chamber: Insights from the Juelich Ozone Sonde Intercomparison Experiment (JOSIE), J. Geophys. Res. 112, D19306, doi:10.1029/2006JD007308, 2007.

Solomon, S., Haskins, J., Ivy, D. J. and Min, F.: Fundamental differences between Arctic and Antarctic ozone depletion, PNAS 2014 111 (17) 6220-6225, doi:10.1073/pnas.1319307111, 2014.

Stavrakou,T., First space-based derivation of the global atmospheric methanol fluxes, Atm. Chem. Phys., 11, 4873-4898, 2013.

Strahan, S.E., A.R. Douglass, and P.A. Newman, The contributions of chemistry and transport to low arctic ozone in March 2011 derived from Aura MLS observations, J. Geophys. Res. Atmos., 118, 1563–1576, doi:10.1002/jgrd.50181, 2013.

Taha, G.; Jaross, G. R.; Bhartia, P. K.: Validation of OMPS LP Ozone Profiles Version 2.0 with MLS, Ozone Sondes and Lidar Measurements, American Geophysical Union, Fall Meeting 2014, abstract #A33J-3322, 2014.

Taylor, K.E.: Summarizing multiple aspects of model performance in a single diagram. J.Geophys. Res., 106, 7183-7192, 2001.

van der A, R. J. , M. A. F. Allaart, and H. J. Eskes, Multi sensor reanalysis of total ozone, Atmos. Chem. Phys., 10, 11277–11294, doi:10.5194/acp-10-11277-2010, www.atmos- chem-phys.net/10/11277/2010/, 2010

van der A, R., M. Allaart, H. Eskes, K. Lefever, Validation report of the MACC 30-year multi-sensor reanalysis of ozone columns Period 1979-2008, MACC-II report, Jan 2013, MACCII_VAL_DEL_D_83.3_OzoneMSRv1_20130130.docx/pdf.

van der A, R. J., Allaart, M. A. F., and Eskes, H. J.: Extended and refined multi sensor reanalysis of total ozone for the period 1970–2012, Atmos. Meas. Tech., 8, 3021-3035, doi:10.5194/amt-8-3021-2015, 2015.

*Vrekoussis, M., Wittrock, F., Richter, A., and Burrows, J. P.: GOME-2 observations of oxygenated VOCs: what can we learn from the ratio glyoxal to formaldehyde on a global scale?, Atmos. Chem. Phys., **10**, 10145-10160, doi:10.5194/acp-10-10145-2010, 2010*

Wennberg, P. O., Mui, W., Wunch, D., Kort, E. A., Blake, D. R., Atlas, E. L., Santoni, G. W., Wofsy, S. C., Diskin, G. S., Jeong, S., and Fischer, M. L.: On the sources of methane to the Los Angeles atmosphere, Environ. Sci. Technol., 46, 9282–9289, https://doi.org/10.1021/es301138y, 2012

Wittrock, F., A. Richter, H. Oetjen, J. P. Burrows, M. Kanakidou, S. Myriokefalitakis, R. Volkamer, S. Beirle, U. Platt, and T. Wagner, Simultaneous global observations of glyoxal and formaldehyde from space, Geophys. Res. Lett., 33, L16804, doi:10.1029/2006GL026310, 2006

WMO (2010), Guidelines for the Measurement of Atmospheric Carbon Monoxide, GAW Report No. 192, World Meteorological Organization, Geneva, Switzerland, 2010.

WMO (2013), Guidelines for the Continuous Measurements of Ozone in the Troposphere, GAW Report No. 209, World Meteorological Organization, Geneva, Switzerland, 2013.

Wunch, D., Wennberg, P. O., Toon, G. C., Keppel-Aleks, G., and Yavin, Y. G.: Emissions of greenhouse gases from a North American megacity, Geophys. Res. Lett., 36, 1–5, https://doi.org/10.1029/2009GL039825, 2009.

Annex 1: Acknowledgements

Listed below are the authors contributing to the sections in this report. The authors contributing to the model description are also provided, as well as acknowledgements to the validation datasets.

Tropospheric reactive gases reactive gases

Annette Wagner, MPG (editor, O3 sondes, GAW data)
Yasmine Bennouna, Valerie Thouret, CNRS-LA (IAGOS)
Harald Flentje, DWD (O3 sondes, GAW data)
Anne Blechschmidt and Andreas Richter, IUB Bremen (GOME-2 NO2, HCHO)
John Kapsomenakis, Christos Zerefos, AA (ESRL)
Natalia Sudarchikova, satellite IR observations (MPG)
Kaj Hansen, Ulas Im, AU (Arctic theme)
Bavo Langerock, BIRA (NDACC)

Tropospheric aerosol

Michael Schulz, MetNo (editor, Aerocom, Aeronet)
Anna Benedictow, Jan Griesfeller, MetNo (Aerocom, Aeronet)
Sara Basart, MTeresa Pay, Oriol Jorba, BSC-CNS (Aeronet, MODIS, AirBase, SDS-WAS NAMEE RC)
Emilio Cuevas, AEMET (Aeronet, MODIS, AirBase, SDS-WAS NAMEE RC)
Harald Flentje, DWD (Backscatter profiles)

Stratospheric reactive gases

Yves Christophe, BIRA (editor, model-satellite intercomparisons)
Simon Chabrillat, BIRA (model intercomparisons)
Annette Wagner, MPI-M (O$_3$ sondes)
Bavo Langerock, BIRA (NDACC FTIR, MWR, UVVIS DOAS, LIDAR)
Anne Blechschmidt and Andreas Richter, IUB-UB Bremen (SCIAMACHY/GOME-2 NO$_2$)

Greenhouse gases

Michel Ramonet, IPSL (ICOS)
Abdelhadi El-Yazidi and Leonard Rivier, LSCE (ICOS)
Thorsten Warneke, UBC (TCCON)
Bavo Langerock, BIRA (TCCON)

Reactive gases and aerosol modeling

Johannes Flemming (ECMWF), Antje Inness (ECMWF), Angela Benedetti (ECMWF), Sebastien Massart (ECMWF), Anna Agusti-Panareda (ECMWF), Johannes Kaiser (KCL/MPIC/ECMWF), Samuel Remy (LMD), Olivier Boucher (LMD), Vincent Huijnen (KNMI), Richard Engelen (ECMWF)

Acknowledgements for the validation datasets used

Hazan, L., J. Tarniewicz, M. Ramonet, O. Laurent and A. Abbaris **(2016)**. *Automatic processing of atmospheric CO$_2$ and CH$_4$ mole fractions at the ICOS Atmosphere Thematic Centre*. Atmospheric Measurement Techniques 9(9): 4719-4736.

Blumenstock, T., F. Hase, M. Schneider, O. E. García, and E. Sepúlveda. 2017. "TCCON data from Izana (ES), Release GGG2014.R1." CaltechDATA. doi:10.14291/tccon.ggg2014.izana01.r1.

De Mazière, M., M. K. Sha, F. Desmet, C. Hermans, F. Scolas, N. Kumps, J.-M. Metzger, V. Duflot, and J.-P. Cammas. 2017. "TCCON data from Réunion Island (RE), Release GGG2014.R1." CaltechDATA. doi:10.14291/tccon.ggg2014.reunion01.r1.

Deutscher, N. M., J. Notholt, J. Messerschmidt, C. Weinzierl, T. Warneke, C. Petri, and P. Grupe. 2017. "TCCON data from Bialystok (PL), Release GGG2014.R1." CaltechDATA. doi:10.14291/tccon.ggg2014.bialystok01.r1/1183984.

Dubey, M. K., B. G. Henderson, D. Green, Z. T. Butterfield, G. Keppel-Aleks, N. T. Allen, J.-F. Blavier, C. M. Roehl, D. Wunch, and R. Lindenmaier. 2017. "TCCON data from Manaus (BR), Release GGG2014.R0." CaltechDATA. doi:10.14291/tccon.ggg2014.manaus01.r0/1149274.

Dubey, M. K., R. Lindenmaier, B. G. Henderson, D. Green, N. T. Allen, C. M. Roehl, J.-F. Blavier, et al. 2017. "TCCON data from Four Corners (US), Release GGG2014.R0." CaltechDATA. doi:10.14291/tccon.ggg2014.fourcorners01.r0/1149272.

Feist, D. G., S. G. Arnold, N. John, and M. C. Geibel. 2017. "TCCON data from Ascension Island (SH), Release GGG2014.R0." CaltechDATA. doi:10.14291/tccon.ggg2014.ascension01.r0/1149285.

Goo, T.-Y., Y.-S. Oh, and V. A. Velazco. 2017. "TCCON data from Anmeyondo (KR), Release GGG2014.R0." CaltechDATA. doi:10.14291/tccon.ggg2014.anmeyondo01.r0/1149284.

Griffith, D. W. T., N. M. Deutscher, V. A. Velazco, P. O. Wennberg, Y. Yavin, G. Keppel-Aleks, R. A. Washenfelder, et al. 2017. "TCCON data from Darwin (AU), Release GGG2014.R0." CaltechDATA. doi:10.14291/tccon.ggg2014.darwin01.r0/1149290.

Griffith, D. W. T., V. A. Velazco, N. M. Deutscher, C. Paton-Walsh, N. B. Jones, S. R. Wilson, R. C. Macatangay, G. C. Kettlewell, R. R. Buchholz, and M. O. Riggenbach. 2017. "TCCON data from Wollongong (AU), Release GGG2014.R0." CaltechDATA. doi:10.14291/tccon.ggg2014.wollongong01.r0/1149291.

Hase, F., T. Blumenstock, S. Dohe, J. Groß, and M.ä. Kiel. 2017. "TCCON data from Karlsruhe (DE), Release GGG2014.R1." CaltechDATA. doi:10.14291/tccon.ggg2014.karlsruhe01.r1/1182416.

Iraci, L. T., J. R. Podolske, P. W. Hillyard, C. Roehl, P. O. Wennberg, J.-F. Blavier, J. Landeros, et al. 2017. "TCCON data from Edwards (US), Release GGG2014.R1." CaltechDATA. doi:10.14291/tccon.ggg2014.edwards01.r1/1255068.

—. 2017. "TCCON data from Indianapolis (US), Release GGG2014.R1." CaltechDATA. doi:10.14291/tccon.ggg2014.indianapolis01.r1/1330094.

Kawakami, S., H. Ohyama, K. Arai, H. Okumura, C. Taura, T. Fukamachi, and M. Sakashita. 2017. "TCCON data from Saga (JP), Release GGG2014.R0." CaltechDATA. doi:10.14291/tccon.ggg2014.saga01.r0/1149283.

Kivi, R., P. Heikkinen, and E. Kyrö. 2017. "TCCON data from Sodankylä (FI), Release GGG2014.R0." CaltechDATA. doi:10.14291/tccon.ggg2014.sodankyla01.r0/1149280.

Liu, Cheng, Wei Wang, and Youwen Sun. 2018. "TCCON data from Hefei (PRC), Release GGG2014.R0." CaltechDATA. doi:10.14291/tccon.ggg2014.hefei01.r0.

Morino, I., N. Yokozeki, T. Matsuzaki, and M. Horikawa. 2017. "TCCON data from Rikubetsu (JP), Release GGG2014.R2." CaltechDATA. doi:10.14291/tccon.ggg2014.rikubetsu01.r2.

Morino, I., T. Matsuzaki, and M. Horikawa. 2017. "TCCON data from Tsukuba (JP), 125HR, Release GGG2014.R2." CaltechDATA. doi:10.14291/tccon.ggg2014.tsukuba02.r2.

Morino, Isamu, Voltaire A. Velazco, Akihiro Hori, Osamu Uchino, and David W. T. Griffith. 2018. "TCCON data from Burgos, Ilocos Norte (PH), Release GGG2014.R0." CaltechDATA. doi:10.14291/tccon.ggg2014.burgos01.r0.

Notholt, J., C. Petri, T. Warneke, N. M. Deutscher, M. Palm, M. Buschmann, C. Weinzierl, R. C. Macatangay, and P. Grupe. 2017. "TCCON data from Bremen (DE), Release GGG2014.R0." CaltechDATA. doi:10.14291/tccon.ggg2014.bremen01.r0/1149275.

Notholt, J., T. Warneke, C. Petri, N. M. Deutscher, C. Weinzierl, M. Palm, and M. Buschmann. 2017. "TCCON data from Ny Ålesund, Spitsbergen (NO), Release GGG2014.R0." CaltechDATA. doi:10.14291/tccon.ggg2014.nyalesund01.r0/1149278.

Pollard, David Frank, John Robinson, and Hisako Shiona. 2019. "TCCON data from Lauder (NZ), Release GGG2014.R0." CaltechDATA. doi:10.14291/tccon.ggg2014.lauder03.r0.

Sherlock, V., B. Connor, J. Robinson, H. Shiona, D. Smale, and D. F. Pollard. 2017. "TCCON data from Lauder (NZ), 120HR, Release GGG2014.R0." CaltechDATA. doi:10.14291/tccon.ggg2014.lauder01.r0/1149293.

—. 2017. "TCCON data from Lauder (NZ), 125HR, Release GGG2014.R0." CaltechDATA. doi:10.14291/tccon.ggg2014.lauder02.r0/1149298.

Strong, K., S. Roche, J. E. Franklin, J. Mendonca, E. Lutsch, D. Weaver, P. F. Fogal, J. R. Drummond, R. Batchelor, and R. Lindenmaier. 2018. "TCCON data from Eureka (CA), Release GGG2014.R3." CaltechDATA. doi:10.14291/tccon.ggg2014.eureka01.r3.

Sussmann, R., and M. Rettinger. 2017. "TCCON data from Garmisch (DE), Release GGG2014.R2." CaltechDATA. doi:10.14291/tccon.ggg2014.garmisch01.r2.

—. 2018. "TCCON data from Zugspitze (DE), Release GGG2014.R1." CaltechDATA. doi:10.14291/tccon.ggg2014.zugspitze01.r1.

Té, Y., P. Jeseck, and C. Janssen. 2017. "TCCON data from Paris (FR), Release GGG2014.R0." CaltechDATA. doi:10.14291/tccon.ggg2014.paris01.r0/1149279.

Warneke, T., J. Messerschmidt, J. Notholt, C. Weinzierl, N. M. Deutscher, C. Petri, and P. Grupe. 2017. "TCCON data from Orléans (FR), Release GGG2014.R0." CaltechDATA. doi:10.14291/tccon.ggg2014.orleans01.r0/1149276.

Wennberg, P. O., C. M. Roehl, D. Wunch, G. C. Toon, J.-F. Blavier, R. Washenfelder, G. Keppel-Aleks, N. T. Allen, and J. Ayers. 2017. "TCCON data from Park Falls (US), Release GGG2014.R1." CaltechDATA. doi:10.14291/tccon.ggg2014.parkfalls01.r1.

Wennberg, P. O., C. M. Roehl, J.-F. Blavier, D. Wunch, and N. T. Allen. 2017. "TCCON data from Jet Propulsion Laboratory (US), 2011, Release GGG2014.R1." CaltechDATA. doi:10.14291/tccon.ggg2014.jpl02.r1/1330096.

Wennberg, P. O., D. Wunch, C. M. Roehl, J.-F. Blavier, G. C. Toon, and N. T. Allen. 2017. "TCCON data from Caltech (US), Release GGG2014.R1." CaltechDATA. doi:10.14291/tccon.ggg2014.pasadena01.r1/1182415.

—. 2017. "TCCON data from Lamont (US), Release GGG2014.R1." CaltechDATA. doi:10.14291/tccon.ggg2014.lamont01.r1/1255070.

Wennberg, P. O., D. Wunch, Y. Yavin, G. C. Toon, J.-F. Blavier, N. T. Allen, and G. Keppel-Aleks. 2017. "TCCON data from Jet Propulsion Laboratory (US), 2007, Release GGG2014.R0." CaltechDATA. doi:10.14291/tccon.ggg2014.jpl01.r0/1149163.

Wunch, D., J. Mendonca, O. Colebatch, N. T. Allen, J.-F. Blavier, S. Roche, J. Hedelius, et al. 2017. "TCCON data from East Trout Lake, SK (CA), Release GGG2014.R1." CaltechDATA. doi:10.14291/tccon.ggg2014.easttroutlake01.r1.

Wunch, D., Toon, G. C., Sherlock, V., Deutscher, N. M., Liu, C., Feist, D. G., & Wennberg, P. O. (2015). The Total Carbon Column Observing Network's GGG2014 Data Version. Tech. rep., California Institute of Technology, Pasadena. doi:10.14291/tccon.ggg2014.documentation.R0/1221662

Copernicus Atmosphere Monitoring Service